

Measuring the level of severity in pharmacoeconomic analyses

An empirical approach

Fredrik Arneberg



Master Thesis
Department of Health Management and Health Economics

UNIVERSITY OF OSLO

May 11, 2012

© Fredrik Arneberg

2012

Measuring the level of severity in pharmacoeconomic analyses: an empirical approach

Fredrik Arneberg

<http://www.duo.uio.no/>

Print: Reprosentralen, University of Oslo

ABSTRACT

Background: Economic evaluations are increasingly being used to inform decision makers regarding the allocation of scarce health care resources. In reimbursement applications to the Norwegian Drug Reimbursement Scheme the cost-utility analysis (CUA) is a common feature, expressing the cost-effectiveness of a treatment. Norwegian Medicines Agency (NoMA), which administers the scheme, is interested in supplementing the existing dossier with measurements of the conditions severity level. During their work on revising the pharmacoeconomic guidelines they identified a lack of empirical studies concerning severity measures (NoMA, 2011b).

Objective: To find out how to operationalize severity measurements in a Norwegian context and investigate the implications of the included severity measures; providing a contribution to the discussion of severity level in Norway.

Method: An explorative approach with a systematic literature review to identify different severity measures. Quality-adjusted life expectancy (QALE) based on Scandinavian data is developed to establish a reference class for the included conditions. The data material is reimbursement applications received by NoMA in the period 2009-2011, including 20 cases in the analysis.

Results: Two severity measures, Absolute shortfall (AS) and Proportional shortfall (PS), are included and operationalized. They both make use of Quality-adjusted life year (QALY) utility values, but have a fundamental difference in how the age of the patient is handled. None of them stands out as being clearly superior regarding existing policy and studies concerning the public opinion.

Conclusions: NoMA and other governmental agencies could get useful information to utilize in resource allocation if doing a more consistent evaluation of medical conditions' severity levels. AS favored younger individuals, but the age component was reduced by including discounting in the procedure. PS was more sensitive to conditions with an expected death in near future independent of patient's age. The thesis identifies various components that have to be addressed to developers of pharmacoeconomic models to achieve valid measurements. This type of methodology are not widespread and it would be interesting to see more research done on applying severity measures to real-life data to develop new and refine existing principles.

Disclaimer: The findings, interpretations and conclusions expressed in the paper are entirely those of the author, and do not represent the views of the Norwegian Medicines Agency.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my supervisor Professor Tor Iversen, at the Department of Health Management and Health Economics, at the University of Oslo, for being available and patience and giving good advices. To the staff and my fellow students at the department: thanks for broaden my horizon and all the fun. A special thanks to Birthe Neset for being the glue keeping it all together.

I would also like to express my gratitude to the Department of Reimbursement at the Norwegian Medicines Agency (NoMA) for giving me the amazing opportunity for this learning process, and the staff for an inclusive attitude. In particular, I want to thank my co-supervisor Morten Aaserud for his dedication to the topic, enthusiasm, patience, and assistance throughout my work.

The topic creates excitement in the health economic environment, and there is a list of additional people I want to thank for their contributions along this journey of personal and professional learning.

Finally, most of all I am grateful for the unconditioned support from my wife Gunnhild. Without your help, none of this would have been possible.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	v
LIST OF TABLES	x
LIST OF FIGURES.....	xi
ABBREVIATIONS AND ACRONYMS	xii
1. INTRODUCTION	1
1.1 Thesis structure.....	2
1.2 QALY and QALE.....	2
2 INSTITUTIONAL FRAMEWORK	4
2.1 Norwegian drug reimbursement scheme	4
2.2 Guidelines and norms concerning severity in the Norwegian health-sector	5
3 THEORETICAL FRAMEWORK.....	6
3.1 Health, disease and severity	6
3.2 Severity.....	6
3.2.1 The severity of illness approach.....	6
3.2.2 Fair innings.....	8
3.2.3 Prospective health	8
3.2.4 QALYs and severity.....	8
4 METHODOLOGY AND DATA.....	10
4.1 Study design	10
4.2 Search strategy.....	10
4.3 Inclusion of measures	11
4.4 Measures of the level of severity	12
4.4.1 Absolute shortfall (AS)	12
4.4.2 Proportional shortfall (PS)	14
4.5 Quality Adjusted Life Expectancy (QALE).....	16

4.5.1	Health-Related Quality of Life (HRQL)	17
4.5.2	Life expectancy	18
4.4.3	Implementing QALE	19
4.6	Data sources and limitations	21
4.6.1	Time horizon and adjustments of the pharmacoeconomic models	22
4.7	Data.....	23
4.7.1	Osteoporosis (Aclasta)	23
4.7.2	Actinic keratosis (AK) (Aldara)	23
4.7.3	Acute coronary syndromes (ACS) (Brilique).....	24
4.7.4	Alzheimer dementia, moderate to severe (Ebixa)	25
4.7.5	Acute Coronary Syndromes (ACS) (Efient)	25
4.7.6	Hormone-dependent metastatic prostate cancer (Firmagon)	26
4.7.7	Chronic myelogenous leukemia (CML) (Glivec)	26
4.7.8	Gastrointestinal Stromal Tumors (GISTs) (Glivec)	26
4.7.9	Non-small-cell lung cancer (NSCLC) (Iressa).....	27
4.7.10	Mild to moderate Ulcerative colitis (UC) (Mezavant)	27
4.7.11	Atrial fibrillation (Multaq)	28
4.7.12	Hepatocellular carcinoma (HCC) (Nexavar)	29
4.7.13	Chronic obstructive pulmonary disease (COPD) (Onbrez).....	29
4.7.14	Osteoporosis (Prolia).....	30
4.7.15	Opioid-induced constipation in cancer treatment (Relistor)	30
4.7.16	Relapse/refractory multiple myeloma (RRMM) (Revlimid)	31
4.7.17	Severe pain/opioid induced constipation (Targiniq)	31
4.7.18	Multiple myeloma (MM) (Thalidomide Pharmion).....	31
4.7.19	Severe chronic hand eczema (CHE) (Toctino)	32
4.7.20	Diabetes type 2 (Victoza)	32
5	RESULTS	34

5.1	Analysis of the operationalization of severity	34
5.2	Ranking of conditions.....	37
6	DISCUSSION	38
6.1	Study objectives.....	38
6.2	Operationalization of severity level.....	38
6.2.1	Consistency challenges when deriving QALYs and QALE	38
6.2.2	Applications focusing on particular age groups or stages of a condition.....	40
6.2.3	Patient populations with a high standard deviation in age	41
6.2.4	Discounting of the level of severity	41
6.3	Implications of the severity measures	42
6.3.1	The valuation of severity by the public	42
6.3.2	Preventive treatments	43
6.3.3	Historically severe conditions	44
6.3.4	Implications of the severity measure procedures	46
6.4	Trade-offs	49
6.4.1	Double counting	52
6.5	Limitations.....	53
6.5.1	EQ-5D values used in the QALE procedure	53
7	CONCLUSION.....	55
	REFERENCES.....	57
	APPENDIX	62
i.	Assessed reimbursement applications.....	62
ii.	EQ-5D utility values.....	63
iii.	Severity calculation	64

LIST OF TABLES

Table 1 Search strategy: literature search in MEDLINE and EMBASE	10
Table 2 Example of using AS in severity level measurement.....	13
Table 3 Example of using PS in severity level measurement.	15
Table 4 Extract from life table	18
Table 5 Inclusion process of reimbursement applications	21
Table 6 Comparison of EQ-5D norms	26
Table 7 the approach for establishing patient's utility values (source: (NoMA, 2012a)	28
Table 8 Conditions sorted after Absolute shortfall (AS).....	35
Table 9 The effect of discounting on the AS scale	47
Table 10 How age influences the PS principle	47
Table 11 Trade-off between cost-effectiveness (ICER).....	50
Table 12 Incremental cost-effectiveness ratio (ICER).....	50
Table 13 Trade-off between cost-effectiveness (ICER).....	51
Table 14 Overview of the assessed applications.....	62
Table 15 Swedish EQ-5D values	63
Table 16 Severity calculations	64

LIST OF FIGURES

Figure 1 Illustration of the construction of Quality Adjusted Life Expectancy (QALE)	16
Figure 2 Illustration of the relationship between HRQL, life expectancy and QALE, and the effect of discounting.....	20
Figure 3 Illustration of the stages of COPD	22
Figure 4 Histogram – showing the distribution of data for the two severity measures.....	34
Figure 5 Scatter plot illustrating the distribution of conditions between the two measures	36
Figure 6 Matrix of a measure of statistical dependence between AS and PS	37
Figure 7 Simplified model of the relationship between disease states.....	43
Figure 8 Trend in mortality from ischemic heart disease.....	45
Figure 9 Variation in the range of AS	47
Figure 10 Fair innings age-weighting	48

ABBREVIATIONS AND ACRONYMS

AS	Absolute shortfall
CUA	Cost utility analysis
DRG	Diagnose related groups
ICER	Incremental cost effectiveness ratio
HDIR	Directorate of Health
HRQL	Health-Related Quality of Life
MAU	Multi-attribute utility instrument
MoH	Ministry of Health and Care Services
NOK	Norwegian Kroner
NoMA	Norwegian Medicines Agency
PS	Proportional shortfall
SSB	Statistics Norway
PTO	Person trade-off
TTO	Time trade-off
QALE	Quality adjusted life expectancy
QALE _{pat}	QALE for the diseased population with the existing treatment
QALE _{pop}	QALE for the general population
QALY	Quality adjusted life years
WTP	Willingness-to-pay

1. INTRODUCTION

Economic evaluations are increasingly being used to inform decision makers regarding the allocation of scarce health care resources. The aim of pharmacoeconomics is to compare the value of one drug or therapy against another when both costs and effect are taken into consideration. These kinds of studies serve to guide an optimal allocation of healthcare resources, in a standardized and scientifically grounded manner. In the reimbursement applications to the Norwegian Drug Reimbursement Scheme (Blåreseptordningen) the cost-utility analysis (CUA) is a common feature, usually expressing the benefits of a drug in quality-adjusted-life-year (QALY). Politicians and other decision makers might want to value different aspects of health when distributing health care resources and a recurring issue is the consideration of the severity of a condition. In Norwegian healthcare policy, a condition's level of severity is one of three core values that are considered when assigning priority; in addition to the efficacy and cost-effectiveness of the intervention (Morland et al., 2010, MoH, 2001).

How should we measure and quantify the severity of a medical condition? Variables as risk of death, loss of physical and physiological functionality, pain, discomfort, and former experienced health are all affecting how people would characterize the level of severity. There exist various methods and principles of measuring a condition's severity within and between medical areas. At Norwegian policy level, there are no accepted standards in how this measurement should be operationalized and consequently governmental bodies as the Norwegian Medicines Agency (NoMA), are using an arbitrary approach.

NoMA acknowledges this shortcoming and experienced a lack of empirical studies in how to operationalize the level of severity when they recently revised the pharmacoeconomic guidelines for the Drug Reimbursement Scheme (NoMA, 2011b). Among the efforts to gain more knowledge about the topic, NoMA made a master thesis proposal that this paper is responding to. By using data from previous reimbursement applications, NoMA wants more knowledge about implications of severity measures and how different disease areas are evaluated. This thesis will investigate different approaches in how to improve the shortcomings that exists, and examine different factors that enables or complicates the implementation of measurement methods.

The research question that is examined in this thesis is:

How can severity measurements be operationalized in applications to the Norwegian Drug Reimbursement Scheme, and what are the implications of alternative measures?

The investigation of the research question is carried out with an explorative approach with a desire to give an empirical contribution to the ongoing debate concerning severity measures.

1.1 Thesis structure

Chapter 1 is the introduction to the thesis and a brief explanation of two central concepts of the thesis. Chapter 2 introduces the institutional context of the research question and tries to define some of the existing Norwegian policy concerning severity level in health economic analysis. Chapter 3 gives a theoretical background presenting philosophical and methodological approaches to the severity issue. Chapter 4 outlines the study design, the search strategy used to identify relevant severity measures and the data material. This chapter takes also a deeper tour into the specific severity principles and the empirical model developed for the analysis. To get a more comprehensive understanding of the data material, a brief introduction is given to the different conditions included. The results of the analysis is presented in chapter 5. In chapter 6, the findings of the analyses are discussed emphasizing the issues of operationalization and implications of severity measures in pharmacoeconomic analyses. Since the thesis also serves an exploring purpose to the topic, attention are given to the requirements and challenges to achieve measurements of validity. Chapter 7 concludes this study and gives suggestions for future research.

1.2 QALY and QALE

Quality-adjusted life year (QALY) have increasingly been used in the assessment of health interventions for the last three decades. QALY is a measure that combines health related quality of life (HRQL) (h) on a scale of 0 to 1, with a time dimension (t), and are expressed with the following equation:

$$(1) \text{ QALY} = h \times t$$

It is assumed that a year of life lived in a state of less than perfect health is worth less than 1, e.g. $0.8 (h) \times 1 (t) = 0.8 \text{ QALYs}$. QALY are often used in cost-utility analyses (CUA) to calculate the ratio of costs to QALYs gained for a particular healthcare intervention. An

advantage of this method is that it makes possible doing comparisons between different interventions for various diseases and conditions (Drummond et al., 2005, Sassi, 2006).

A related principle to QALY is the Quality Adjusted Life Expectancy (QALE), which is the number of years of perfect health one can expect to live from a certain age, and it is a summary measure of mortality and HRQL across different stages of life. QALE as a measure summarizes the overall health for the population and provides comparisons among different groups (Jia et al., 2012). The concepts and the underlying factors are being discussed throughout the thesis.

2 INSTITUTIONAL FRAMEWORK

2.1 Norwegian drug reimbursement scheme

The Norwegian Medicines Agency (NoMA) is responsible for preparing recommendations and when relevant passes resolutions concerning acceptance of drugs to the reimbursement scheme (Blåreseptordningen). NoMA assesses annually approximately 110 reimbursement cases that include applications for reimbursement of generic drugs, new administration methods, new strength, new active substances and new indications for existing substances. The applications must include a pharmacoeconomic analysis of the drug according to the Guidelines for Pharmacoeconomic Evaluations (NoMA, 2012b). Based on reimbursement applications submitted, as well as NoMAs own reports; health outcomes, cost effectiveness, and the overall budget impacts are evaluated. The perspective in the cost effectiveness analysis is a societal, with some limitations. NoMA will come to a decision on its own, or make a recommendation to the Ministry of Health and Care Services (abbreviated: MoH). MoH assesses applications that have an estimated net budget impact for the National Insurance' pharmaceutical budget (Folketrygden) of more than five million NOK in the fifth year of reimbursement. In these cases the reimbursement decision is taken by the Storting (The Parliament) after a proposal from the Ministry (MoH, 2012, NoMA, 2005)

Comprehensive reimbursement applications include a cost-utility analysis (CUA). The advantage of this well-established method is that it makes it possible to compare the value of funding for entirely different competing programs. Research indicates that priority setting between programs according to cost per QALY methods in many cases does not correspond well with the societal preference of what is fair and ethical (Nord, 1993a, Ubel, 1999). The QALY approach is based on an idea to maximize health production, but a society such as the Norwegian, also seeks to prioritize according to degree of severity (where effective treatment does exist) (NoMA, 2005, NOU, 1997). A drug can be accepted for reimbursement if the criteria stated in § 14–13 in the medicine regulation are fulfilled:

- a) treatment or prevention of a severe condition
- b) long-lasting condition
- c) clinical effectiveness (efficacy)
- d) cost-effective

(NoMA, 2011a)

2.2 Guidelines and norms concerning severity in the Norwegian health-sector

Two reports from the Lønning Commissions have provided terms of concepts for development of health care policy in Norway, leading for instance to the Regulations on the Prioritization of Health Services and the Right to Health Care and the Patient Rights Act of 2001 (MoH, 2001, NOU, 1997). One of the important elements of this work is the recommendation that priority should be given according to three criteria:

- Severity of the condition
- The expected outcome from the intervention
- A reasonable cost-effectiveness ratio

The commission has defined that severity could be an expression of lost lifetime and quality of life compared to the average population, if treatment is not initiated. According to the commission's opinion a condition could be determined to be severe based on an assessment of:

- Death risk or loss of functionality
- The extent of physical and physiological loss of functionality
- Pain, physical or physiological discomfort (e.g. short-windedness, dizziness, anxiety)

The commission has not been more specific, and operationalizing of the topic is therefore open to interpretation for both governmental agencies, different organizations and the patients themselves. To establish a goal of equal health state for everyone is characterized as unrealistic. However, conditions that deviates the most from normal health should be given the highest priority. The severity level of a condition itself does not indicate a high priority if the patient group only can expect a marginal or no outcome of the treatment (NOU, 1997)

It is a common comprehension that young people's health is a precious resource in a society, and some would claim that this is an integrated concern in what people understand as a severe condition. How the age of the patient should influence a measure of severity is not determined by the policymakers. The age dimension is not emphasized in the severity definitions from the Lønning Commission, but the Guidelines for Priority Setting is stating that a reduction of life expectancy is relevant for the right to receive necessary healthcare (MoH, 2001).

3 THEORETICAL FRAMEWORK

There are different ways to approach the severity concept. In this chapter, I will present different theories, statements and principles concerning severity.

3.1 Health, disease and severity

As a fundamental approach to severity, Christopher Boorse's influential (1977) definition of health and disease as theoretical concepts is chosen:

1. The *reference class* is a natural class of organisms of uniform functional design; specifically, an age group of sex of a species.
2. A *normal function* of a part of process within members of a reference class is a statistically typical contribution by it to their individual survival and reproduction.
3. A *disease* is a type of internal state that is either an impairment of normal functional ability, i.e. a reduction of one or more functional ability, i.e. a reduction of one or more functional abilities below typical efficiency, or a limitation on functional ability caused by environmental agents.
4. *Health* is the absence of disease (Boorse, 1977)

The magnitude of deviation from what are expected as *health* could therefore be a definition of the level of severity of a condition. This also seems to be in accordance with the more specified definitions from the Lønning Commission (NOU, 1997) and the two approaches serves as a theoretical basis in this thesis.

3.2 Severity

To be able to have a distinct approach to the topic it is of importance to be able to refine the concept and to distinguish it from related concepts such as equity, fairness and equality.

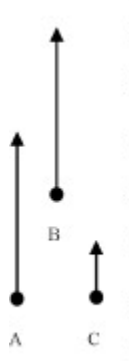
In the following sub-chapters, different theories of severity will be presented. The theories stress different factors in the determination of how to compute and understand severity level.

3.2.1 The severity of illness approach

The severity of illness approach embodies the feeling that people with severe conditions (e.g. facing immediate death) must be prioritized. For instance, if individual A can be taken from 0.4 to 0.6 on a 0-1 utility interval scale, and individual B can be taken from 0.6 to 0.8, then

the former improvement would be valued higher than the latter, all else being equal. A more comprehensive example has been illustrated by Shah (2009), adapted from Nord (1993b):

	<u>Problem level</u>	<u>Example (in terms of mobility)</u>
	1 None (healthy)	
	2 Slight	Can move about anywhere, but has difficulties with walking more than 2km.
	3 Moderate	Can move about without difficulty at home, but has difficulties in stairs and outdoors.
	4 Considerate	Moves about with difficulty at home. Needs assistance in stairs and outdoors.
	5 Severe	Can sit. Needs help to move about – both at home and outdoors.
	6 Very severe	To some degree bedridden. Can sit in a chair part of the day if helped up by others.
	7 Completely disabled	Permanently bedridden.
	8 Dead	



The arrows represent possible gains from intervention for individual A, B and C. A and B has the same capacity to benefit from intervention (three-level improvement in health state), but in the absence of treatment, A finds himself in a worse health state than B – he is more severely ill than B. The severity of illness approach suggests that other things being equal, the health gains accruing to A should be valued more highly than those accruing to B since there is an intrinsic value in helping the worst-off. Those who advocate allocation according to severity would maybe prefer to give C priority over B even though the expected health gain from treatment is smaller. Thus, the severity approach can involve a sacrifice of aggregate health in order to target the worst off (Shah, 2009)

There is not only severity at the time of intervention that is taken into account in this principle, but also expected severity in future years in case of non-intervention. A problem with this approach may be that substantial differences in health prospects may exist not only because of different illnesses, but also because of age differences. Another counterintuitive situation arises when relatively healthy individuals must wait until their health deteriorates sufficiently for them to satisfy the severity of illness criterion (Nord, 2005, Shah, 2009, Richardson et al., 2011, Stolk et al., 2004).

Another definition of severity is launched by Hansson et al. (1994) who states that the severity of disease can be defined as prognosis measured in QALYs, i.e. expected remaining life years adjusted for the quality of life during these years, with the existing treatment method. This definition is in contrast with the mentioned definition above and is more influenced by the time dimension and preference for younger individuals.

3.2.2 Fair innings

The equity principle of fair innings reflects the feeling that everyone is entitled to some “normal” span of health. Anyone failing to achieve this has been cheated, whilst anyone getting more than this should consider any additional years a sort of bonus. Fair innings is characterized by being a notion of equity being outcome based. It is about lifetime experience, not about an individual’s state at any particular point in time. It reflects an aversion to inequality, and it is quantifiable and even in common parlance it has strong numerical connotations. Death at 35 is viewed very differently from death at 90, and age at death is the key variable which is most often focused on (Williams, 1997). In an allocation situation with scarce health care resources, priority is given to the individuals that are most far away from achieving their “innings”. Unlike other approaches, fair innings includes the retrospective aspect of health.

3.2.3 Prospective health

In contrast, this approach bases its severity or equity emphasizing on the expected ill health (including death) over the remaining lifetime if left untreated. The principle focus on caring for those with poor health prospects that corresponds with that people may wish to devote considerable resources to improve the health of seriously ill people, particularly those who are in risk of facing immediate death. It also incorporates a fairness concern when it contributes to equalize the distribution of health. Efforts and priority should be distributed initially to those who can expect the worst prospective health if they are left untreated.

3.2.4 QALYs and severity

In CUA, the focus is how much gain in QALYs an intervention cause. If QALYs are used to determine the severity of a condition, the decision maker needs to have knowledge about the originating of the values. The QALY maximization rule entails distributive neutrality, which means that it does not incorporate concerns for how benefits are distributed across individuals. A “QALY is a QALY is a QALY” under all circumstances, regardless of to whom they accrue and the context in which they are enjoyed. So priority is given to the patient group that can gain the most from treatment independent of age and discomfort. However, people may be willing to sacrifice aggregate health gain in order to direct resources towards those who are worst off in terms of the severity of their initial condition (Olsen, 1997). Studies also show that the health maximization solution does not have a very strong support among people (Cuadras-Morato et al., 2001)

A discussed issue is whether severity and societal concerns is actually captured in the initial personal utility assessments of health states used in priority settings, which is used in the development of QALY-values. Some support that this is captured, while others argue that the respondents have never expressed any opinion about priority setting or society value, but a quantification of the disutility they would feel with different states of illness. Since the respondents does not have a societal perspective we might excluding this kind of fairness reflections. To remove this imbalance there has been proposed various approaches to establish equity weights to measure and integrate society's aversion to inequality in health outcomes (Nord, 1999). On the other hand, it is said that the QALY maximizing rule includes two central equity characteristics:

- A QALY gain has the same societal value irrespective of who is generating it (independent of income, race, gender)
- QALY maximizing is in accordance with the utilitarian theory, which means that the proper course of action is the one that maximizes the overall "happiness" (utility)

Obviously this rule does not include all equity considerations, and it is demonstrated that there are different understandings of equity across populations (Olsen, 2006). Some societies, like the Norwegian, tend to emphasize egalitarianism rather than the utilitarian view (Nord, 1999). Egalitarianism is a philosophy that supports equality between individuals. However, by modifying the QALY approach one might end up with a less transparent and easily understood process.

It is challenging to design health care decision-making systems that consider all stakeholders' perspectives. From a health economic perspective it is of importance to have consistent systems, while other groups may emphasize flexibility in severity valuations. The latter is illustrated in a report concerning QALYs and severity of illness from The Citizen Council which is an advisory body of The National Institute for Health and Clinical Excellence (NICE) in the UK (NICE, 2008). Here they state that they think it is of importance to take the severity of a disease into account when making decisions. Furthermore, they say that they would prefer NHS (National Health Service) to take severity "into consideration" alongside the cost and clinical effectiveness evidence rather than including severity in the calculation of QALY. The reasons given is that including severity in the calculation of QALY will lessen the transparency of a committee's decision (NICE, 2008).

4 METHODOLOGY AND DATA

4.1 Study design

This study is based on a descriptive platform where the aim is to provide information about how measures of medical condition's severity can be operationalized and included in pharmacoeconomic analysis. Numerical data concerning medical conditions and pharmacoeconomic models are derived from reimbursement applications assessed by NoMA in the period 1.1.2009-31.12-2011. These data have been linked with life expectancy data from Statistics Norway (SSB) to create measures that are in touch with reality. There is in addition derived disease specific information to be able to present and discuss the measurements in an appropriate context.

4.2 Search strategy

To be able to identify different measurement principles of severity it is necessary to perform a literature search. The method used is inspired by the handbook "Doing a systematic review" (freely translated) from The Norwegian Knowledge Centre for the Health Services (NOKC, 2011). I have performed a literature search in MEDLINE and EMBASE to get a comprehensive search, and to assure that I capture the variety of medical journals and other relevant sources. The two databases are using different methods to index articles and different keywords are being used in the search algorithms to cover the area of interest. The important concept of the search is to include articles that deal with severity of illness in a health care prioritization context. To improve the quality of the search there has been used different relevant articles to check if the search algorithm captured them.

MEDLINE		EMBASE	
Searches	Results	Searches	Results
exp "health care economics and organizations"/	1104599	exp quality adjusted life year/	8364
exp "Severity of Illness Index"/	138767	exp disease severity/ or severity of illness.mp.	956751
Health Priorities/	7902	exp economic evaluation/	178489
1 and 2 and 3	56	1 and 2 and 3	756

Table 1 Search strategy: literature search in MEDLINE and EMBASE

To include various subheadings and to ensure a broad search, most keywords had to be expanded (exp). After duplicate removal in EndNote, the total number of articles was 762. Most of these were focusing on specific diseases and the number of relevant articles that was included was 14. Few of these articles are precisely matching the research question, but they are a good basis for locating references through a snowball method. By going through the reference lists and looking for citing of other publications, a diverse amount of sources was acquired.

4.3 Inclusion of measures

The focus of the analysis is the severity of a condition. It is of importance to be specific so the measurement do not get influenced by other concerns such as equity, costs, burden of disease for society etc. An ideal method is a universal measure that captures the severity of a condition across different disease areas (making it comparable), in an unbiased, accurate and comprehensive manner, and in a transparent way. Different measures of severity were identified during the search process and the most relevant are:

- Absolute shortfall (HDIR, 2011, Lee et al., 2010, Jia et al., 2012)
- Before and after treatment (Hagino et al., 1998)
- Disability-adjusted life year (DALY) (Drummond et al., 2005)
- Disease specific measures (Revicki and Wood, 1998)
- DRG-based severity measures (Romano and Chan, 2000)
- Proportional shortfall (Stolk et al., 2004)
- Value number (FHI, 2011)

Most of the identified methods are disease specific with a usage in a clinical setting, and cannot be used as a generic instrument on a system level. Others are focusing on the severity of the pre- and post-intervention health states and the numbers of persons treated. Both the before/after-measure and the severity-adjusted DRGs approach are treatment-oriented, and do not indicate the condition's severity. The DALY measure is developed by WHO and the World Bank, and is commonly used in low- and middle-income countries to assess the magnitude of disease, health risks, and premature death. As a measure of disease burden DALY could seem to be an appropriate tool in severity measurements; but it has some characteristics that makes it incompatible with the objectives of the Drug Reimbursement Scheme:

- Age weights that give lower weight to years of the young and the elderly
- Cannot include impacts of side-effects and co-morbidities
- Unable to distinguish between alternative healthcare interventions
- Has only seven health states in addition to dead and healthy making it difficult to catch marginal differences
- The weighting used in the measurements are not preference based, but based on valuations by a panel of healthcare workers
- Low validity and reliability

(Drummond et al., 2005, Fox-Rushby and Cairns, 2005)

Another opportunity which is suggested by the Norwegian Institute of Public Health (FHI, 2011) and referred to as “Value number” (freely translated) is to simply subtract the QALY value for the condition from 1, e.g. individuals affected by condition X has a QALY value of 0.6, and thereby have a severity value of 0.4 (1-0.6). The problem with this method is that the data material in the reimbursement applications only reports the QALY gain for a condition with an existing and with a new treatment, and do not emphasize where on the 0-1 scale the improvement finds place. The procedure is therefore not applicable.

Absolute and Proportional shortfall seems to be the measures that are best transferable to the current topic. Both are QALY-based in estimating health utilities, they have a transparent calculation of severity levels, and will be introduced in the two next sections. Following by a section where the development of a quality-adjusted life expectancy (QALE) model is explained.

4.4 Measures of the level of severity

4.4.1 Absolute shortfall (AS)

The procedure was first identified when doing explorative research for this thesis in a discussion document written by the Directorate of Health (HDIR, 2011). The procedure has no exact name, so for simplicity I will name it Absolute shortfall (AS). It is easy to understand and have a practical approach, which is illustrated in a study conducted by Lee et al. (2010), where the expected loss-of-QALE due to stroke are estimated. The procedure is identified in different variants for a variety of issues, e.g. in burden of disease studies (Muennig et al., 2006). It assumes that the point of departure of measuring severity should be the life

expectancy of the general population and the HRQL of the age group that the patient belong to. The measure is defined as:

$$(2) AS = QALE_{pop} - QALE_{pat}$$

Eq. (2) defines AS as the QALE of the general population ($QALE_{pop}$) minus the QALE of a particular patient group (with the existing treatment) ($QALE_{pat}$). The size of the minuend will therefore be decisive in the valuation of severity level and consequently younger individuals will have a more favorable starting point than the elderly will. For the Norwegian population the QALE of a newborn is 67 (undiscounted) and represents the maximum value of the measure. In the following table, three age groups with three different diseases that reduce QALE with 40, 60 and 90 percent are being measured to illustrate the fundamental mechanisms of the principle:

Patient's age	30	50	70
QALE for age group	50	30	10
Disease reduces QALE with	40 %	60 %	90 %
Absolute shortfall	20	18	9
Ranking	1	2	3

Table 2 Example of using AS in severity level measurement.

The patient being the one with the largest number of lost QALE is determined to be the worst off in respect of severity. How old the patient is, or the average age of a patient group will therefore heavily affect the level of severity according to this principle, and may be an issue of criticism. The objections against the QALY methodology are in some cases also applicable to this procedure. The discrimination of the elderly is the most relevant objection being exemplified in Table 2 where the eldest that also has the condition with the highest proportion of QALE-reduction is ranked as the third most severe (Petrou and Renton, 1993). The severity definition stated by Hansson et al. (1994) seems to be compatible with the mechanisms and outcomes of AS.

In a utilitarian approach to the CUA the aim is to prioritize the interventions that produce the most QALYs (Dolan et al., 2005). With AS we are also able to rank the objects after their score, but in this thinking, it is all about the condition and the patient group that experience the highest loss of QALE. A crucial difference between the two principles is that CUA is intervention oriented, whilst AS is condition oriented. AS's impact on the decision material

would therefore lessen the effect of the QALY maximization rule, which is one of the essentials of including severity into health economic analysis.

Egalitarianism advocates equal shares of the good being distributed. In this case, it can be understood as equal access to health care across individuals. It can also be comprehended as equalizing health within a population. AS has a foundation in the average individual of the population since QALE are being used as reference group when computing level of severity. The mechanisms of AS could therefore contribute to more equality. It is important to emphasize that this reasoning do not include the approach of strong egalitarianism, but a milder specific egalitarianism (Olsen, 1997, Olsen et al., 2003).

Parts of the fair innings argument may be included as a characteristic of this principle. As described in chapter 3.2.2, the argument takes the view that there is some span of years that we consider a reasonable life. Anyone who does not achieve this span of years has been short-changed, and should be given priority, and this may result in an age-based rationing policy (Rivlin, 2000). Fair innings has total life span as focus meaning the closer you are to your “limit”, the less priority is given. The same type of logic is recognizable for AS since it includes the dimension of time in the severity measure.

4.4.2 Proportional shortfall (PS)

The principle of Proportional shortfall was introduced by the Dutch researcher Elly Stolk in 2004 (Stolk et al., 2004) and has reached a broad consensus of use in the Netherlands (van de Wetering et al., 2011). It is though still considered as an invention within the field of health economic evaluation, and is not extensively used elsewhere. The authors encourages efforts like the analysis conducted in this thesis, when they claim that the ultimate test of the practical value of the concept is to apply it to real-life situations (Stolk et al., 2004). Proportional shortfall (PS) is stated by the authors to be a procedure that is characterized by being an intermediate position between prospective health and the fair innings approach (van de Wetering et al., 2011). As fair innings, PS is concerned with disease-related QALE-loss. At the same time, in accordance with prospective health, it takes the remaining QALE expectation with the existing treatment into account. It assumes that measurement of severity in health should concentrate on the fraction of QALYs that people lose relative to their life expectancy, and not on the absolute number of QALYs lost or gained (van de Wetering et al.,

2011). PS is measured on a scale from 0 (no loss of health) to 1 (maximum loss of health, or immediate death) using the following fraction:

$$(3) PS = \frac{AS}{QALE_{pop}} = \frac{QALE_{pop} - QALE_{pat}}{QALE_{pop}} = 1 - \frac{QALE_{pat}}{QALE_{pop}}$$

The numerator reflects the loss of QALE due to disease (i.e. AS), as described in the previous section, and the denominator reflects the remaining life expectancy, which is HRQL-adjusted in this analysis by using QALE. The patient with the largest proportion of lost QALE is thereby determined to be the one worst off in terms of severity. Thus, it wishes to equalize patients within their own scope of potential health, and therefore it differs from fair innings and prospective health in the sense that it compares individuals in relative terms to determine who is worse off. As a result, this principle does not discriminate elderly, but it is more sensitive for patients with a short life expectancy (Stolk et al., 2004, van de Wetering et al., 2011).

A characteristic of Proportional shortfall is that all patients who face a threat of immediate death will score 1, irrespective of their age, and will all receive equal weight (horizontal age weighting). Likewise, if a young patient with an expectation of 40 QALEs loses 20 QALY, he or she will get the same severity weight as an older patient with 2 QALEs loses 1 QALY: both patients lose 50 % of their remaining life expectancy (Stolk et al., 2004, van de Wetering et al., 2011).

Table 3 illustrates the mechanisms of PS. The same age groups and the same three diseases that exemplified in the AS section. The ranking made my PS is the opposite of AS' and illustrates the fundamental difference between the principles.

Patient's age	30	50	70
QALE for age group	50	30	10
Disease reduces QALE with	40 %	60 %	90 %
QALE lost (AS)	20	18	9
Proportional shortfall	40 %	60 %	90 %
Ranking	3	2	1

Table 3 Example of using PS in severity level measurement.

4.5 Quality Adjusted Life Expectancy (QALE)

Most of the severity approaches and measures compare the loss of health to the patient group or the general population. AS and PS use the general population. The reimbursement applications are designed for the Norwegian market, and it is reasonable to use the health of the general Norwegian population as a comparator. Both Lønning (p. 5) and Boorse (p. 6) emphasize the principle of doing comparisons with the population the patient belongs. Thus, a more reasonable comparison and measurement than if for example a global standard are being used, which is the case for the DALY measure (Drummond et al., 2005).

Quality Adjusted Life Expectancy (QALE) is the number of years of perfect health one can expect to live from a certain age, and is a summary measure of mortality and Health-Related Quality of Life (HRQL) across different stages of life. QALE as a measure summarizes the overall health for the population and provides comparisons among different groups, and may be useful to guide policies in healthcare issues. HRQL assesses an individual's perception of quality of life in relation to health or in response to the onset of a disease (Preedy and Watson, 2010). Life expectancy is a summary of the mortality rates of age-specific groups in the population. Since HRQL differs across different stages of life, calculating an adjusted life expectancy gives a more complete measure for assessing overall health (Jia et al., 2011).

QALE may be a contribution to making the analysis more precise than not adjusting the loss of health for differences in HRQL and life expectancy. The structure of QALE calculation is quite complicated and an attempt to illustrate the structure is done in Figure 1.

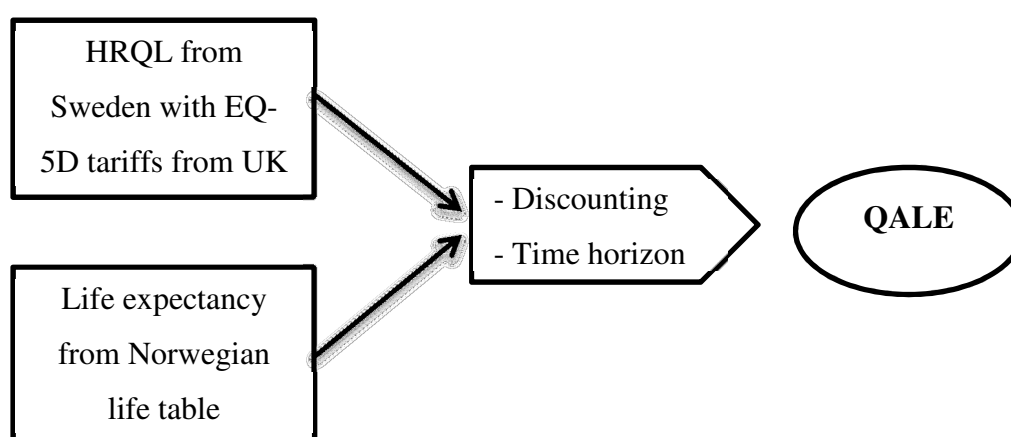


Figure 1 Illustration of the construction of Quality Adjusted Life Expectancy (QALE)

4.5.1 Health-Related Quality of Life (HRQL)

There has not yet been developed HRQL-scores based on Norwegian data that is useful for the purpose of this thesis. It is therefore necessary to use data from close to similar populations. Among the Scandinavian countries, Burström and Rehnberg's (2006) report concerning the HRQL of the population in Stockholm County, Sweden, seems to provide the most appropriate values to use in the severity calculations. To be able to quantify the HRQL data the commonly used EQ-5D measure was implemented in their study. EQ-5D is a standardized instrument for use as a measure of health outcome. It is applicable to a wide range of health conditions and treatments, and it provides a descriptive profile and a single index value for health status (EuroQol, 2012). The EQ-5D requires respondents to classify their own health status according to five health dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) using three levels of severity (no problems, moderate problems, and severe problems) and defines 243 (3^5) different health states (Drummond et al., 2005). In general, EQ-5D is the measure used in the reimbursement applications to derive QALY-values.

The measurements were included in a public health survey in 2002, and the results are based on data from more than 30 000 individuals aged between 18 and 84 (see appendix ii for more details). The survey was a cross-sectional self-administered postal questionnaire based on a stratified study design, where gender and area (25 municipalities and 18 districts in Stockholm city) were used as stratum variables (Burström and Rehnberg, 2006).

There is no EQ-5D preference weights developed in Sweden, so the authors have used British preference weights (tariffs) based on the Time-trade-off (TTO) technique in the EQ-5D index score. Preference weights determine how much each of the five dimensions should count. In TTO the respondents are asked to choose between remaining in a state of ill health for a period of time, or being restored to perfect health but having a shorter life expectancy. Despite the great extensiveness of EQ-5D it is not automatically superior to other measures (Brazier and Ratcliffe, 2007). The implications will be addressed in the limitations chapter.

The Swedish study sample consists of people aged 18 to 84. To get a model that incorporates a total life span of utility values it is necessary to supplement with values for years <18 and >84. A pragmatic approach was chosen to make a lifelong QALE structure:

- The average utility value for years 18-40 is 0.85. Children and adolescents in the age interval 0-18 are assumed to have a relatively high HQRL, so the value for these years has been set to 0.9.
- From age interval 75-79 to 80-84, the utility value decreases with 0.05 (6.6 %). The same level of deterioration is applied on the age intervals 85-89 and 90-105.

Next step in the construction of QALE is to link HRQL with life expectancy. A life table from Statistics Norway (SSB) contains the necessary data and is used as source to model life expectancy (Table 4) (SSB, 2011).

4.5.2 Life expectancy

5 Life tables¹, 2010

Age x	Survivors at age x			Death at age x to x+1			Life expectancy - remaining years at age x			Probability of death at age x, per 1 000 (Ungraduated)		
	lx			dx			ex			qx		
	Both sex ²	Males	Females	Both sex ²	Males	Females	Both sex ²	Males	Females	Both sex ²	Males	Females
0	100 000	100 000	100 000	276	305	246	81.04	78.85	83.15	2.763	3.053	2.457
1	99 724	99 695	99 754	27	16	40	80.26	78.09	82.35	0.273	0.156	0.397
2	99 696	99 679	99 715	5	10	0	79.28	77.11	81.39	0.049	0.095	0.000
3	99 692	99 670	99 715	11	13	10	78.29	76.11	80.39	0.115	0.128	0.101
4	99 680	99 657	99 705	5	3	7	77.30	75.12	79.40	0.050	0.032	0.068
5	99 675	99 654	99 698	8	7	10	76.30	74.13	78.40	0.084	0.066	0.103
6	99 667	99 647	99 688	7	10	3	75.31	73.13	77.41	0.067	0.098	0.034
7	99 660	99 637	99 684	10	7	14	74.31	72.14	76.41	0.102	0.067	0.138
8	99 650	99 631	99 670	8	13	3	73.32	71.14	75.42	0.085	0.133	0.035
9	99 642	99 617	99 667	13	10	17	72.33	70.15	74.42	0.132	0.096	0.168
10	99 628	99 608	99 650	8	13	3	71.34	69.16	73.44	0.080	0.126	0.033
11	99 620	99 595	99 647	3	3	3	70.34	68.17	72.44	0.032	0.031	0.033
12	99 617	99 592	99 644	8	16	0	69.34	67.17	71.44	0.080	0.156	0.000
13	99 609	99 577	99 644	14	15	13	68.35	66.18	70.44	0.141	0.152	0.129
14	99 595	99 561	99 631	14	24	3	67.36	65.19	69.45	0.140	0.241	0.032
15	99 581	99 537	99 628	14	12	16	66.37	64.20	68.45	0.141	0.121	0.161

Table 4 Extract from life table made by Statistics Norway (SSB, 2011). Range: 0-105 years.

The life table is based on the probability of dying between two age levels (number of years) and is indicated with the symbol q_x in which x stands for years. The point of departure is a selected cohort of 100 000 persons of both genders, in which survivors at age x is expressed by l_x :

$$(4) l_x = l_{x-1}(1 - q_{x-1})$$

In which $(1 - q_{x-1})$ is the probability that a person aged $x-1$ will be alive one year later. The number of deceased persons by their exact age x to $x+1$ year in the selected cohort is:

$$(5) d_x = l_x \cdot q_x$$

To calculate the life expectancy at age x we need the auxiliary aggregates contemporary survivors and total remaining lifetime. The contemporary survivors of age x ,

$$(6) L_x$$

is the average number of persons of age x . If deaths are assumed to fall evenly throughout the year, contemporary survivors are:

$$(7) L_x = \frac{1}{2} (l_x + l_{x+1})$$

Total remaining lifetime at age x is the number of years *contemporary survivor* x -year-olds in the cohort have to live:

$$(8) T_x = L_x + L_{x+1} + \dots + L_{\omega}$$

Which ω is equal to the highest age (105 years). Expectation of life at age x is equal to the total of remaining years of life at age x , divided by the number of survivors at age x .

$$(9) \text{Life expectancy} = e_x^0 = \frac{T_x}{l_x}$$

Using the described formula to calculate life expectancy for a 50-year-old male would be to divide the number of remaining life years by the remaining individuals in the cohort:

$$\frac{2\,953\,497}{95\,992} = 30.77$$

4.4.3 Implementing QALE

Outcome measures over the full cycle of care for a medical condition should account for both survival and HRQL in estimating QALE. QALYs are primarily used to adjust someone's life expectancy based on the levels of HRQL they are predicted to experience throughout their lifetime, or part of it. The number of QALYs lived in one year is simply:

$$(10) \quad \text{QALYs lived in one year} = 1 * \text{HRQL with } Q \leq 1;$$

From this descends that someone's QALE at age a can be defined as:

$$(11) \quad \text{QALE} = \sum_{t=a}^{a+L} \text{HRQL}_t$$

Where L is the residual life expectancy of the individual at age a , and t represents individual years within that life expectancy range

When implementing QALE in the severity measurement it is of importance to use the same approach as the developer of the pharmacoeconomic model included in their analyses. The most decisive components are the health measurement tool (e.g. EQ-5D), time horizon of the model and the discount rate. Attention is given to these concepts throughout the chapter. Incorporating discounting represents a value judgment on preference for current to future consumption (Egan, 2002). A higher discount rate means less weight to future benefits and NoMA recommends a 4 % annually discount rate for both costs and health effects.

The following chart is based upon the different described components in order to clarify the relationships in the above sections. The y-axis need some interpretation to give sense: The life expectancy values are measured on a scale from 0-1 and communicate the probability of becoming one year older for a person being 0 years in 2010 (SSB, 2011). This means for example that if this person lives until he is 20 years old, it is a probability of 99.53 % that the person will become 21 and so on. HRQL is the average HRQL for a Swedish population on a scale from 0-1. QALE is life expectancy HRQL-adjusted. Since life expectancy is changing over time and varies for different age classes, the QALE trend line are interpreted as the quality adjusted life expectancy for a newborn in 2010. The difference between QALE and life expectancy reflects the HRQL people place on living with morbidity versus living that time in perfect health. When discounting QALE we can see that the graph changes shape communicating the “time preference” of valuing future health benefits at a lower rate.

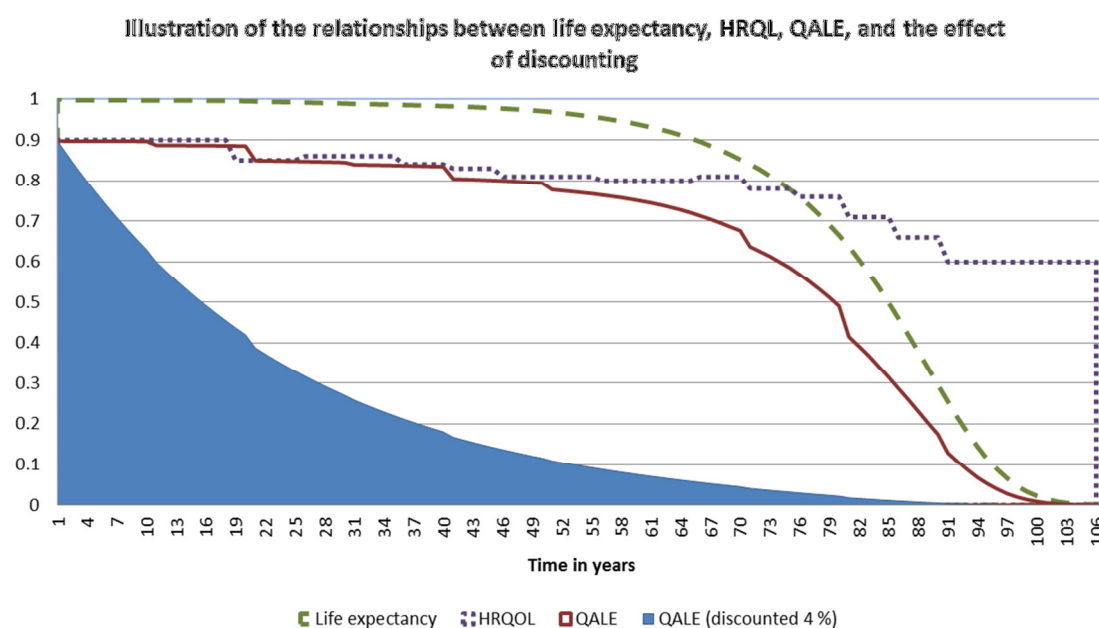


Figure 2 Illustration of the relationship between HRQL, life expectancy and QALE, and the effect of discounting

4.6 Data sources and limitations

Data are obtained from drug reimbursement applications that are downloaded from NoMAs database archive. The content of the dossier depends on the objection of the application and the following inclusion criteria are used for the analysis:

- Cost utility analysis (CUA)
- Data about severity are expressed through QALYs
- Information about the models' time horizons and starting age or median age for the patient population is accessible

To get a manageable volume of the data only applications received and assessed by NoMA during the time period 1.1.2009-31.12.2011 are included in the analysis. Table 5 illustrates the inclusion process communicating the number of applications left in the inclusion in the third column. During this period, 152 reimbursement applications were received by NoMA and of this, 70 were characterized as “long applications”, which means it is a comprehensive dossier containing a CUA. The applicant withdrew three of these and one was rejected by NoMA. The 25 applications that were marked as being under assessment by NoMA were excluded. Applications that have been rejected by NoMA by other reasons than low data quality are included in the analysis.

(1)	Total applications received in period 2009-2011	153
(2)	Total of long applications (CUA)	70
(3)	Excluding applications drawn back/rejected	66
(4)	Excluding applications still being assessed	41
(5)	Excluding applications that didn't fulfill inclusion criteria 2 and 3	21
(6)	Applications included in analysis	20

Table 5 Inclusion process of reimbursement applications

The majority of the long applications that NoMA assesses are cost minimum analysis (CMA) that does not include a pharmacoeconomic model with utility values. The number of relevant applications consequently reduced to 20, which all used EQ-5D scores to derive QALY values. This is positive for the validity of the analysis since the same instrument is used in development the HRQL data used in QALE.

The export of valid data is an intricate and time-consuming task assessing a number of sources to reduce the risk of e.g. using incorrect utility values or wrong median age of the

patient population. The time spent on each case varied from one hour to three days. It is a stepwise process where each source has to be assessed:

1. The reimbursement report from NoMA.
2. The dossier attached to the reimbursement application
3. References used in the application
4. If attached to the application, a health economic model
5. If necessary: find data from other sources which can verify assumptions made in 1-4

4.6.1 Time horizon and adjustments of the pharmacoeconomic models

There are variations between the applicants regarding the time horizon they use in their models, often done to reduce uncertainty in their analysis. To meet this variation two methods are undertaken in the extraction of QALE values. For models with a lifetime perspective, lifelong QALE is used (e.g. osteoporosis). For other conditions, like Chronic obstructive pulmonary disease (COPD), where the treatment regimens only lasts for a certain time, only QALE within the time frame of the treatment is extracted. Figure 3 is an example that illustrates the severity stages of COPD. Drugs are sometimes specialized to treat specific stages of a disease, so to be able to catch the effects of the treatment they choose to focus the models perspective to a certain period. If an intervention has effect on mortality NoMA requires a lifetime perspective (NoMA, 2012b). This topic is discussed more in chapter 6.2.2.

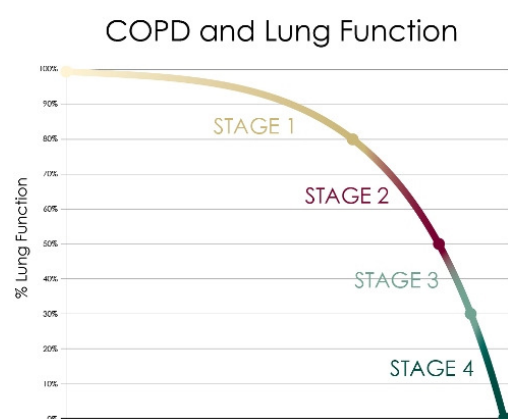


Figure 3 Illustration of the stages of COPD (UHS-PruittCorporation, 2012)

Most of the models included in this analysis are Markov models, which are useful when a decision problem involves risk that is continuous over time, when the timing of events is important, and when important events may happen more than once. In this kind of models it is assumed that the patient is always in one of a finite number of discrete health states, called

Markov states (Sonnenberg and Beck, 1993). Time is handled as a series of discrete cycles, where events of interest are counted either at the beginning or at end of each cycle. In real life, these events can occur at any time within each cycle and are best represented as a continuous probability function. One of the basic assumptions underlying Markov models is that transitions can occur only once in each cycle. In the construction of QALE it is important to facilitate so transitions (mortality) can occur at any time during the cycle to counteract under- or overestimation. In the life table data from SSB this adjustment is done and it is not necessary to do further adjustments (SSB, 2011).

4.7 Data

The 20 reimbursement applications vary in their perspective, time horizon, discounting practices, data sources and underlying condition. To understand the context and the outcome of the severity measures, it is necessary to make a brief introduction to each of the cases. I will briefly explain the nature of the condition (pathophysiology), indication for treatment, and the modeling of the pharmacoeconomic analysis. In cases with distinct characteristics related to the research question, a short discussion will be undertaken. For additional information about the cases see appendix i and iii.

4.7.1 Osteoporosis (Aclasta)

Generic name: Zoledronate.

Osteoporosis is a bone disease that leads to an increased risk of fracture. It does not only cause fractures, but also bedridden secondary complications that might be life threatening in the elderly (WHO, 2003). The model used was initially a Markov model, but was transformed to a “discrete event” model with the advantage that it generates a complete patient profile for each patient and the individual history of fracture, thus a fracture risk can be individually calculated. The time horizon of the model is lifetime perspective and the comparator is no treatment. The applicant has included the same Swedish EQ-5D values and Norwegian life tables as used for QALY calculation in this thesis. This will contribute to a more valid severity evaluation.

4.7.2 Actinic keratosis (AK) (Aldara)

Generic name: Imiquimod.

AK, also known as solar keratosis, presents as scaly lesions on the skin predominantly on chronically sun exposed sites on fair-skinned individuals. AK is a skin-colored, or reddish

brown or yellowish black, ill-defined round or irregularly shaped macule or papule with a dry, firmly adherent scale. The lesion size is generally between 1-3 mm in diameter but can be up to several centimeters. AKs may exist singly in isolation or, more commonly, as multiple and confluent lesions. AK is considered an early stage of cancer and part of a disease continuum that includes squamous cell carcinoma an invasive form of non-melanoma skin cancer (Lucas, 2006). A decision model has been structured as a decision tree. Patients are followed in the model for one year after initial treatment. According to the applicant a one year time horizon is sufficient to capture the majority of relevant costs and consequences attributable to AK treatment, as well as inclusion of the risk of AK recurrence. A complicating factor with the chosen approach is the selection of the value 1 on the 0-1 scale as the health utility for persons with no AK.

Apparently, it seems like the AK patient population has a higher HRQL than the normal population, since it is modeled with a higher score (0.993). The highest score for an age group in QALE is 0.9, which are for nine-year olds. This makes the validity of a severity measurement for this condition low. A simple adjusting may be done by setting $QALE_{pop}$ to 1 that will result in AS: 0.007 and PS: 1 %. This is though a rough adjustment and is not satisfying in a real-life decision making since the uncertainty of the conclusion high. This application could be excluded, but is included since it serves the purpose of illustrating the challenges of operationalizing severity measures.

4.7.3 Acute coronary syndromes (ACS) (Brilique)

Generic name: Ticagrelor.

ACS is an umbrella term for acute conditions that result from a reduction in blood flow to the heart muscle. These conditions range from unstable angina (UA) and non-ST-elevation myocardial infarction (NSTEMI) to ST-elevation MI (STEMI). In UA/NSTEMI a blood clot partly occludes an artery and as a result some of the heart muscle being supplied by the affected artery dies or is at high risk of dying. In STEMI the coronary artery is generally blocked off by a blood clot, and as a result much of the heart muscle being supplied by the affected artery starts to die (Widimsky et al., 2010). The model is a two-part construct with a one-year decision tree, based on data from the PLATO study, and a Markov model for long-term extrapolation to ensure that all major clinical and resource generating events that a patient may experience throughout the course of their remaining life are captured. The model

was run for a period of forty years ensuring that costs and health outcomes were estimated until all patients, in principle, had died.

4.7.4 Alzheimer dementia, moderate to severe (Ebixa)

Generic name: Memantine.

Alzheimer's disease (AD), the most common type of dementia, is a chronic progressive neurodegenerative disorder with an insidious onset. It is characterized by a progressive loss of cognitive and intellectual function, the ability to perform the activities of daily living as well as emergence of behavioral and psychological symptoms. Life expectancy of sufferers with AD is reduced. The median survival period from the diagnosis of Alzheimer disease is approximately 6 years (range 1-16 years), with about one-third of the time being spent in the severe stage (McKhann et al., 1984, Mohs and Cohen, 1988, Geldmacher et al., 2006). The model has a 5-year time horizon. This time frame is considered appropriate as AD represents a chronic progressive disease state with a low life-expectancy among patients with moderate to severe AD. Health utilities of patients were assessed with the EQ-5D, applying a time trade-off (TTO) algorithm. If the patients were unable to answer themselves, next to kin where used for response. It is known that people not affected by a particular disease rate the health problem more severely than people who have the disorders (Gray et al., 2011). In the case of AD, this phenomenon is obviously more uncertain.

4.7.5 Acute Coronary Syndromes (ACS) (Efient)

Generic name: Prasugrel.

ACS is a life threatening condition and refers to any group of clinical symptoms associated with acute myocardial ischemia (MI) with or without infarction. Besides Stroke and Peripheral Arterial Disease (PAD), ACS represents a manifestation of atherosclerosis which is usually precipitated by acute thrombosis, induced by a ruptured or eroded atherosclerotic plaque, with or without concomitant vasoconstriction, causing a sudden and critical reduction in blood flow. Patients with ACS have a high mortality rate, primarily resulting from sudden deaths or the development or recurrence of acute MI (Anderson et al., 2007, Bassand et al., 2007). There is used a Markov model consisting of 15 monthly cycles (representing the trial phase), a single cycle of nine months duration to round the Markov trace to two years, and 38 annual cycles to complete a 40 year model time horizon (lifelong). There is applied an EQ-5D norms from the US which deviates from the Swedish norms used in the QALE calculation in this thesis.

Age	US	Sweden	Deviation
45-54	0.84	0.81	3.6 %
65-74	0.78	0.795	-1.9 %

Table 6 Comparison of EQ-5D norms

In this case, this deviation is only a theoretical problem since the variation is minimal and mortality is the major driver of treatment effect. In a different setting where changes in HRQL were sensitive for changes, the problem could be more critical.

4.7.6 Hormone-dependent metastatic prostate cancer (Firmagon)

Generic name: Degarelix.

Prostate cancer is a form of cancer that develops in the prostate, a gland in the male reproductive system. The cancer cells may metastasize from the prostate to other parts of the body, particularly lymph nodes and the bones. Symptoms may be pain, difficulty in urinating, problems during sexual intercourse, or erectile dysfunction. Approximately 75 % of patients with untreated localized prostate cancer will develop locally progressive disease within 10 years, and 65 % of these will die from cancer. Half of patients with untreated metastatic disease will die within three years (Albertsen et al., 1998). With early detection, treatment can be effective. The model's time-horizon covers period from initiation of hormone therapy to patient death.

4.7.7 Chronic myelogenous leukemia (CML) (Glivec)

Generic name: Imatinib.

CML is a rare, cancer of the white blood cells. It is a form of leukemia characterized by the increased and unregulated growth of predominantly myeloid cells in the bone marrow and the accumulation of these cells in the blood (Faderl et al., 1999). The horizon of the Markov model is lifetime (240 cycles or 60 years) and is expected to capture all relevant costs and effects resulting from the treatment.

4.7.8 Gastrointestinal Stromal Tumors (GISTs) (Glivec)

Generic name: Imatinib.

GISTs are soft tissue sarcomas that arise from primitive mesenchyme cells and are located primarily in the stomach (50%) and small intestine (25%). GISTs are considered rare and

occur in 10-20 per one million people. The indication in the application is for patients with inoperable and/or metastatic malign GISTs. The pharmacoeconomic Markov model has a lifetime perspective, according to the model to be 20 years or approximately 1040 weeks/cycles (i.e. when almost 100 % of the patients in the model are dead). The company has only efficacy data for 225 weeks/cycles, that means that significant extrapolation has been done and will be associated with great uncertainty. Consequently the model was run with a horizon of 10 years (520 weeks/cycles). Since this is a mortal condition and treatment is characterized as life prolonging (median survival time using comparator has been 1.5-3-4 years and Glivec treatment prolongs survival time with 5 years) the QALYs for the condition is compared with lifetime QALE.

4.7.9 Non-small-cell lung cancer (NSCLC) (Iressa)

Generic name: Gefitinib.

Lung cancer is the most common cause of cancer related death. There is still a lack of effective treatments; five-year survival is around 15 percent, and 60-80 percent of patients will die within the first year of diagnosis. If left untreated, this growth can spread beyond the lung in a process called metastasis into nearby tissue and, eventually, into other parts of the body (Branden, 2008). The relevant patient population for treatment with Iressa is patients with advanced or metastatic NSCLC with positive EGFR mutation status in all lines of therapy. A relatively simple Markov model was developed to model disease progression and estimate costs and QALYs over a lifetime horizon, 6 years, with the different treatment strategies under investigation.

4.7.10 Mild to moderate Ulcerative colitis (UC) (Mezavant)

Generic name: Mesalazine.

UC is a chronic disease of the colon, or large intestine that includes characteristic ulcers, or open tiny sores producing pus and mucus. The main symptom of active disease is usually constant diarrhea (sometimes bloody), of gradual onset. Currently, there is no medical cure for ulcerative colitis (TCCFA, 2011). There is used a Markov model to estimate costs and effects and the model has 9 health stages over a period of 5 years. The model's time horizon is not further explained in the application or in the reimbursement report from NoMA.

4.7.11 Atrial fibrillation (Multaq)

Generic name: Dronedarone.

AF is the most common sustained cardiac arrhythmia (irregular heartbeat) in the western world and is a major cause of hospitalization, morbidity and mortality, long-term risk of stroke and congestive heart failure as the leading cause. AF is a complex, progressive and life-threatening disease, with a 1 in 4 life-time risk for individuals over the age of 40 (Pai and Varadarajan, 2007, Go et al., 2001). The patient group is adult clinically stable patients with a history of, or current, non-permanent atrial fibrillation, to prevent recurrence or to lower ventricular rate. A state-transition model with patient-level simulation (discrete event) with a lifetime horizon (28 years) was used in the analysis with a start age for the population at 72.

Factor	Utility	SD	Duration
Constant	1.061	0.045	na
<i>Demographic characteristics</i>			
Age	-0.003	0.001	na
Male sex	0.067	0.008	na
<i>Health states</i>			
AF symptoms	-0.084	0.001	1 month*
Stroke	-0.295	0.036	Lifetime
CHF	-0.190	0.014	Lifetime
ACS	-0.145	0.047	1 year

*Lifetime if patient develops permanent AF.

Table 7 the approach for establishing patient's utility values (source: (NoMA, 2012a))

The utility weights used, are based on a transformation from utilities in a heart study to EQ-5D values. According to the table, one should subtract 0.003 from the constant for each year to get an age-adjusted HRQL. This principle deviates from the regular EQ-5D procedure and may be a source of bias in a severity evaluation. A gender-neutral QALE (as used in this thesis) may increase the difference even more, and in this application, there were close to a 50-50 distribution between genders. A comparison of the health utility value of a non-diseased 72 year old has been done to prove how considerable this type of calculations might deviate from each other:

- Determination of a healthy year as a 72 year old in the Multaq application
 - o $1.061 + (-0.003 \times 72) = 0.845$
- Determination of a healthy year as a 72 year old using QALE methodology (HRQL*probability of death)
 - o $0.77 \times 0.98093 = 0.756$
- Difference: $(0.845 - 0.756) / 0.845 = 11 \%$
- Difference if gender-neutral QALE = $((0.845 + (0.067/2)) - 0.756) / 0.8785 = 14 \%$

This works as an illustration, but it is necessary to comment that the applicant might have used other values if they were required to measure the level of severity.

4.7.12 Hepatocellular carcinoma (HCC) (Nexavar)

Generic name: Sorafenib.

HCC is the most frequent type of liver cancer, being responsible for about 90 % of primary malignant liver tumors in adults. It is the fifth most common cancer worldwide and the third most common cause of cancer death. HCC is associated with a dire prognosis with median survival following diagnosis reported as less than 1 year and for advanced HCC, less than six months (Llovet et al., 2003). In order to account for the progressive and evolutionary nature of HCC the applicant has used a Markov model. Time horizon for the cost-effectiveness analysis is a patient's life time, with the possibility of analyzing for trial duration or up to patient lifetime in annual increments (e.g. 1 year, 2 years , 3 years, et c . up to 20 years). After 14 years less than 1 % of patients are still alive and thus for the purposes of reporting 14 years is considered a lifetime horizon. The application was rejected due to several identified factors that made the validity of the results questionable, with the extrapolation of efficacy data as the absolute most important. A standard deviation of 10.7 years in median patient age could have a significant effect in an evaluation of the level of severity.

4.7.13 Chronic obstructive pulmonary disease (COPD) (Onbrez)

Generic name: Indacaterol.

Chronic obstructive pulmonary disease (COPD) is the co-occurrence of chronic bronchitis and emphysema, a pair of commonly co-existing diseases of the lungs in which the airways become narrowed, resulting in a gradual loss of lung function. The symptoms of COPD range from chronic cough, sputum production, and wheezing to more severe symptoms, such as dyspnea, poor exercise tolerance, and signs or symptoms of right-sided heart failure. COPD develops slowly. Symptoms often worsen over time and can limit the patient's ability to do activities of daily life (nhlbi, 2012). The health economic model, a Markov cohort model, is based on a recognized structure. This model has an unusual 3 year time frame with 12 weeks cycles. The applicant defends the choice of time horizon with that a longer time horizon will considerably increase the uncertainty, since most of patients shift to second and third line treatment and thereby complicates the model's patient flow. The patient group is characterized by having a stage 1 (mild) COPD. It is therefore necessary to value the QALE of the general population within the same time frame as used in the application (age 64-67) which will rate

the condition less severe than if a COPD patient cohort was followed throughout their lifetime.

4.7.14 Osteoporosis (Prolia)

Generic name: Denosumab.

Osteoporosis is asymptomatic until the severity of disease manifest with the occurrence of fractures, particularly fractures of the spine, hip and forearm. It is characterized by low bone mineral density, which is a measure of bone strength. Patients having a low bone mineral density have a significantly increased risk of fractures (Hagen G, 2010). Besides the negative impact on the quality of life of the individual, osteoporosis is also a very costly disease for society. The societal costs associated with the disease are also expected to increase in the future due to the increasing incidence of fractures, in addition to increased life expectancy (Gauthier et al., 2011). A significance of this condition is the high treatment costs, and a slightly improvement in utility results in cost-efficiency. The cycle length in the model is 6 months and all patients are individually followed through the model from the age of treatment initiation to their time of death (or an age of 100 years). The patient population is females with mean age 71. The utility values are based on a Swedish study where the utility values for the general population differ from the utility values used in this thesis (Lundberg et al., 1999).

4.7.15 Opioid-induced constipation in cancer treatment (Relistor)

Generic name: Methylnaltrexone.

Constipation is recognized as a condition with slow and incomplete bowel evacuation and a pathological increase in the digestive tract transit time. It is followed by symptoms such as a bloated stomach, discomfort, pain and pressure feeling, and has in general a negative influence on the quality of life of cancer patients. The use of opioids is a frequent cause of non-obstructive constipation. The relevant patient group is characterized by advanced cancer illness receiving palliative care (Movik E, 2009). The Markov model is developed by The Norwegian Knowledge Centre for the Health Services and are not compatible with a regular assessment. The cost-effectiveness of the treatment has therefore partly been based on valid arguments from the applicant. There was no data on life expectancy or horizon for the model, but taken the severity of the pathophysiology into consideration it is assumed a lifetime horizon.

4.7.16 Relapse/refractory multiple myeloma (RRMM) (Revlimid)

Generic name: Lenalidomide.

Multiple myeloma, also known as plasma cell myeloma, is a fatal progressive cancer characterized by an excessive number of abnormal plasma cells in the bone marrow and the overproduction of intact monoclonal immunoglobulin. Without treatment, survival is around 6 to 12 months after the diagnosis. The median survival time is around 3 years with conventional chemotherapy (patients' ≥ 65 years) (Rajkumar and Kyle, 2005). The pharmacoeconomic model is a discrete event simulation (DES) that can address specific, real settings of care for patients with RRMM. Although Markov models are the most commonly used technique for pharmacoeconomic analysis today, the approach does not have the flexibility required to respond to the individual variations in time to response, duration of response, and survival that are typical among patients with RRMM. The time horizon of the model is lifetime without emphasizing the exact number of years.

4.7.17 Severe pain/opioid induced constipation (Targiniq)

Generic name: Oxycodone/Naloxone.

This condition concerns patients with severe pain that only can be adequately managed with opioid analgesics, in accordance with the medically approved indication. Furthermore, the application for reimbursed use of Targiniq is limited to palliative care of patients within the final stages of life. This implies that terminally ill cancer patients with severe pain that only can be managed with opioid analgesics will constitute the patient population in question for reimbursement. The target patient population consists of terminally ill patients with limited life expectancy. Patients with incurable cancer require qualified treatment for alleviation of pain during the last months of their lives. The duration of the palliative phase for these patients will however vary. Base case time horizon was set to 6 months and included in a cost-utility model developed in Microsoft Excel. Since the time horizon of the model is limited to 3-12 months, discounting of costs and outcomes is not required.

4.7.18 Multiple myeloma (MM) (Thalidomide Pharmion)

Generic name: Thalidomide.

MM is a fatal progressive cancer affecting bone marrow plasma cell development. It is estimated that MM accounts for 1 % of all malignant diseases and 10 % of all blood cancers. Age, gender and race significantly affect the incidence of MM. Age is the most significant

risk factor: the median age at diagnosis is between 63 and 70 years and only 2% of patients are younger than 40 (Durie, 2001). The median age in the analysis is 79 and will therefore be a typical example of different valuation by AS and PS. Without treatment, survival is six to 12 months after the diagnosis. The median survival time is approximately:

- Three years with conventional chemotherapy, in non-transplant-eligible patients aged 65 and older
- Five years with intensive chemotherapy combined with haematopoietic ASCT in transplant-eligible patients under the age of 65 (Sirohi and Powles, 2004)

A Markov model with lifetime horizon is adopted for the analysis.

4.7.19 Severe chronic hand eczema (CHE) (Toctino)

Generic name: Alitretinoin.

Hand eczema is considered chronic if it has persisted for 6 months or more in the absence of external causative factors. Additionally, patients may be classified as suffering from severe, refractory CHE if there is no disease improvement or transient improvement with potent corticosteroid treatment given for a span of 4 weeks, and when patient have avoided known allergens or irritants and used skin emollients and protective measures. CHE is the leading occupational disease and contributes to sick leave, compensation, early retirement and the domestic, social, and psychological impact of hand eczema is considerable (Fowler, 2008). The Markov model used in the reimbursement application follows a theoretical cohort of 1000 severe CHE patients over a number of treatment iterations, for approximately 22 years, allowing for changes in health states at the end of each treatment cycle. A mapping exercise from DLQI (The Dermatology Life Quality Index) to EQ-5D was done based on published literature. The application was rejected by NoMA based on the reason that the submitted data and analyzes provided too small and uncertain evidence on efficacy and cost-effectiveness compared to existing treatment.

4.7.20 Diabetes type 2 (Victoza)

Generic name: Liraglutide.

Diabetes type 2 is a metabolic disorder that is characterized by high blood glucose in the context of insulin resistance and relative insulin deficiency. Due to its high prevalence, it is one of the largest health care challenges in the world. As well as having a high prevalence and increasing incidence, the condition is associated with significant morbidity and mortality

(Wild et al., 2004). The base case is for a population with BMIs reaching into the high 20s. The model is based on a series of inter-dependent sub-models that simulate the complications of diabetes (angina, myocardial infarction, congestive heart failure, stroke, peripheral vascular disease, diabetic retinopathy, macular edema, cataract, hypoglycemia, ketoacidosis, lactic acidosis, nephropathy and end-stage renal disease, neuropathy, foot ulcer and amputation, and non-specific mortality). Each sub-model in the pharmacoeconomic analyses has a semi-Markov structure and uses time, state, time-in-state and diabetes type dependent probabilities derived from published sources. The time horizon was set to patient lifetime (40 years).

5 RESULTS

AS has a positively skewed distribution with a mean value of 4.45 indicating that AS have relatively few high, very severe, values. PS has a bimodal distribution with two distinct peaks indicating that the majority of observations are divided between the minimum and maximum of the scale. The mean value is 0.46, which is the middle of the scale ranging from 0-1. Both measures have a relative high standard deviation.

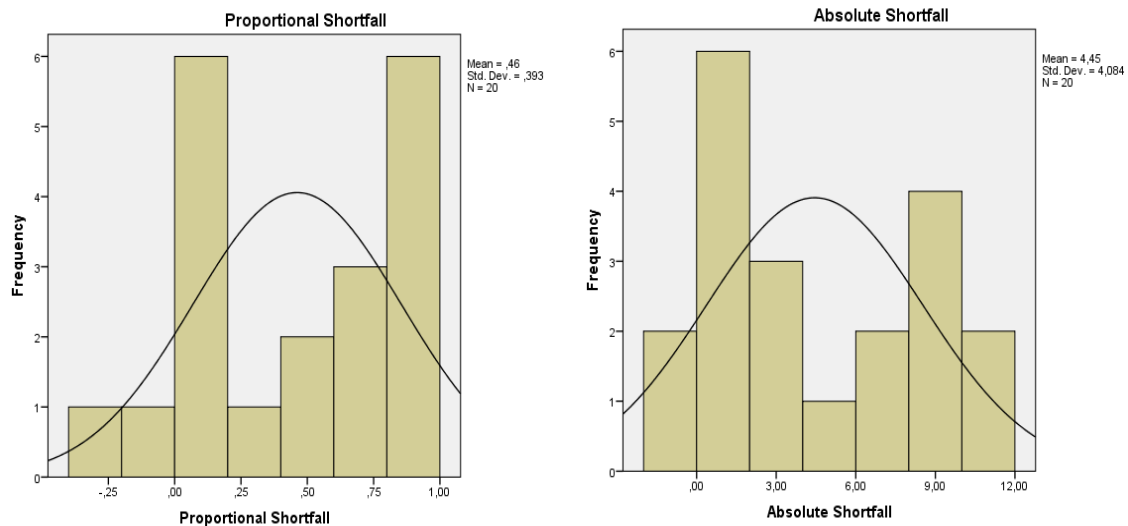


Figure 4 Histogram – showing the distribution of data for the two severity measures

5.1 Analysis of the operationalization of severity

The results are communicated through a detailed table (Table 8) and through a scatter plot and will be discussed more in the following chapter. Out of 20 cases, the Aldara case distinguishes with low comparable validity of the estimates (as described in 4.7.2). However, it serves a benefit of exemplifying the challenging assignment of accurately measuring and comparing the level of severity in pharmacoeconomic analysis.

The first column of the succeeding table presents the trade name and the indication of the treatment. In addition, a specification of the type of treatment has been stated as an abbreviation (see underneath the table for explanation). Column two to six (*d, a, b, c, t*) illustrates the data that are included in the calculation of the severity values in the two last columns. The number in the time horizon-column (*t*) is the length of the model used in the reimbursement application, indicating either the patient age the model are run until (marked ■), or the time span independent of patient age. Some applicants state that they have a lifetime perspective, but does not inform about the specific time horizon. The QALY values in column

a are calculated by the pharmaceutical companies and are not necessary endorsed by NoMA.

The table is ranked by AS. Ranking by PS is displayed in brackets in the last column.

Product name and condition	Discount rate	QALYs established treatment	QALE (general pop.)	Median age /start age in model	Time horizon	AS	PS
	d	a	b^*	c	T	$b-a$	$(b-a)/b$
Targiniq - Severe pain opioid induced constipation (P)	0 %	0.21	11.76	59	0.5	11.55 (1)	98 % (1)
Iressa - Non-small-cell metastatic lung cancer (P)	4 %	0.87	12.34	57	Lifetime (6)	11.47 (2)	93 % (2)
Glivec (GIST) - Gastrointestinal Stromal Tumors (P)	4 %	2.17	11.47	60	Lifetime (10)	9.30 (3)	81 % (6)
Glivec (CML) - Chronic myelogenous leukemia (P)	4 %	5.00	13.89	51	Lifetime (60)	8.89 (4)	64 % (9)
Nexavar - Hepatocellular carciNoMA (P)	2.5 %	0.69	9.22	67	Lifetime (14)	8.53 (5)	93 % (3)
Revlimid - Relapse refractory multiple myeloma (P)	4 %	2.19	10.54	63	Lifetime	8.35 (6)	79 % (7)
Relistor - Opioid-induced constipation in cancer treatment (P)	0 %	0.71	7.48	72	Lifetime (0.33)	6.77 (7)	91 % (4)
Firmagon - Hormone-dependent metastatic prostate cancer (P)	4 %	0.97	7.13	73	Lifetime	6.16 (8)	86 % (5)
Victoza - Diabetes type 2 obese patients (RS)	4 %	7.18	12.34	57	Lifetime (40)	5.16 (9)	42 % (11)
Ebixa - Alzheimer dementia (RS)	3 %	1.21	4.82	80	Lifetime (5)	3.61 (10)	75 % (8)
Thalidomide Pharmion - Multiple myeloma (P)	4 %	2.29	4.99	79	Lifetime (29)	2.70 (11)	54 % (10)
Toctino - Severe chronic hand eczema (RS)	2.5 %	11.55	13.62	48	22	2.08 (12)	15 % (13)
Efient - Acute Coronary Syndromes (SP)	4 %	9.74	11.16	61	Lifetime (40)	1.42 (13)	13 % (14)
Aclasta - Osteoporosis (SP)	4 %	6.17	7.13	73	Lifetime	0.96 (14)	13 % (15)
Multaq - Atrial fibrillation (SP)	4 %	6.54	7.48	72	Lifetime (28)	0.93 (15)	13 % (16)
Onbrez - Chronic obstructive pulmonary disease (COPD) (RS)	4 %	2.14	2.96	64	3	0.82 (16)	28 % (12)
Mezavant - Ulcerative colitis (IR)	3 %	3.44	3.85	43	5	0.41 (17)	11 % (17)
Brilique - Acute coronary syndromes (SP)	3 %	9.68	9.97	67	Lifetime (40)	0.28 (18)	3 % (18)
Prolia - Osteoporosis (SP)	4 %	8.05	7.57	71	Lifetime (100) ■	-0.20 (19)	-3 % (19)
Aldara - Actinic keratosis (SP)	0 %	0.99	0.79	69	1	-0.20 (20)	-25 % (20)

Table 8 Conditions sorted after Absolute shortfall (AS). * = QALE is discounted at the same rate as used in the pharmacoeconomic models developed by the applicant, reported in column d, and the procedure involves d,c and t. IR = induction of remission; P = palliative treatment; RS = relieving of symptoms and/or maintenance treatment; SP = secondary prevention. See the appendix for a more comprehensive table. ■ = Model is run until patient age is 100.

The disease areas that are being rated as most severe is clear when studying the results of AS.

Cancer conditions of various types hold the eight first rows. Six out of these have a median

age lower than 70 being a contributor to the discrepancy between AS and PS in the ranking of the upper part of the table. Both the table and the chart on the next page illustrate the concurrence and variation between the methods. To get consistent values QALE has been discounted at the same rate as used for QALYs in the reimbursement applications. This is less time-consuming than modifying the complicated models used in the reports (not all are available), which also could become a source of error if not done in collaboration with the developer. Not discounting both components (columns named a and b) at the same rate would lead to inconsistencies and impossible comparisons and conclusions. The differing cases of Targiniq, Nexavar and Relistor have a lower or no discount rate at all, because the conditions are associated with a near-term death. To get a more correct severity level the $QALE_{pop}$ component of the severity calculation has been discounted with 4 %. We can do this since $QALE_{pat}$ is so low and would probably not be noticeable reduced by discounting.

A feature of the illustrations is how age influences the level of severity computed with AS. Patient groups with a low median age will consequently be rated more severe with AS. PS will on the other hand rate the conditions more severe if the expected lifetime with disease is short.

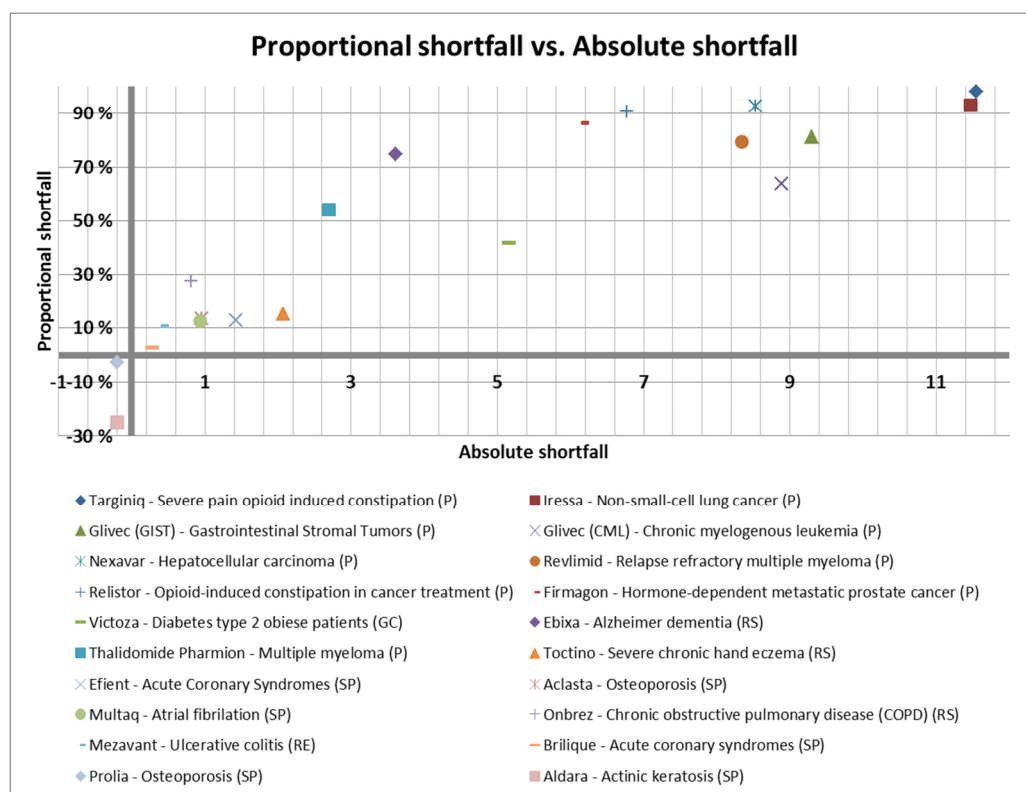


Figure 5 Scatter plot illustrating the distribution of conditions between the two measures. IR = induction of remission; P = palliative treatment; RS = relieving of symptoms and/or maintenance treatment; SP = secondary prevention

5.2 Ranking of conditions

In Table 8 there are differences between the AS and PS measures if the conditions are ranked by their severity. The differences and possible patterns will be discussed in the following chapter. There was conducted a nonparametric statistical measurement of AS and PS to investigate the strength and direction that exist between the two methods. Spearman's rank order correlation coefficient that is defined as the Pearson correlation coefficient between the ranked variables was chosen since the data do not have a normal distribution. The determined relationship between the two severity measurements of 20 conditions was a strong, positive correlation, which was statistically significant at a 0.01 level ($r_s(18) = .933$, $P = .000$).

Correlations			Proportional Shortfall	Absolute Shortfall
Spearman's rho	Proportional Shortfall	Correlation Coefficient	1,000	,933 ^{**}
		Sig. (2-tailed)	.	,000
		N	20	20
	Absolute Shortfall	Correlation Coefficient	,933 ^{**}	1,000
		Sig. (2-tailed)	,000	.
		N	20	20

^{**}. Correlation is significant at the 0.01 level (2-tailed).

Figure 6 Matrix of a measure of statistical dependence between AS and PS

The statistical measurement indicates that the two methods have corresponding rankings of conditions, but it does not provide information of a more comprehensive character. Interpreted, this means that the ranking differences are minor between the measures, exemplified by the measures having close to the same conditions on the upper and lower part of the ranked table.

6 DISCUSSION

6.1 Study objectives

In general, the analysis in this thesis intends to shed light on the issue of valuing severity in pharmacoeconomic evaluations used to assess whether or not drugs should be granted general reimbursement in Norway. The composition of this chapter is designed to explain findings in a manner that makes it useful as input in the ongoing severity discussion in Norway.

The objective of the thesis is divided into two areas of interest: operationalization and implications. In this regard, two methods for measuring the level of severity were chosen and constructed to evaluate the severity level of conditions assessed by NoMA in the period 2009-2011 for inclusion or not in the Norwegian drug reimbursement scheme. These two measures were also recommended to be used in the proposed pharmacoeconomic guidelines for the drug reimbursement scheme, but were withdrawn in the final edition due to uncertainty around the impacts.

To my knowledge, this type of empirical approach has never been conducted in Norway and similar research design is not identified through the search strategy. Therefore an explorative and descriptive research design is applicable. In this chapter, the relationship between Norwegian policy on severity and the outcomes of the measures are discussed. The shortcomings of the methodologies and the challenges concerning the data material are also central.

6.2 Operationalization of severity level

6.2.1 Consistency challenges when deriving QALYs and QALE

There are some fundamental consistency challenges in the methodology used to derive QALY and QALE, and therefore affects the results of AS and PS. Different sources for the two numbers used in the severity measures ($QALE_{pop}$ and $QALE_{pat}$) could establish biases in the analysis. Most pharmacoeconomic analysis is developed abroad, and is modified to the Norwegian context by the pharmaceutical companies. I would like to highlight the four most occurred challenges which could cause bias:

1. QALYs are computed with different methods and tools in the cases

- In this thesis, Swedish EQ-5D values and Norwegian life expectancy are important components of the analysis. Only the Aclasta case (Osteoporosis) has used the same data.
2. Diverse sources of acquiring utility values
 - For some conditions, the researchers had interviewed patients with first-hand experience, while others used the members of the public as respondents. A number of empirical studies have indicated that the patients themselves tend to place higher values on dysfunctional health states than the members of the general population who do not have similar experience (Brazier and Ratcliffe, 2007). Other applicants have used the mapping solution that enables health state utility values to be predicted for use in CUA when no preference-based measure has been included in the study (Brazier et al., 2010).
 3. A healthy individual in the applicants analysis has a higher utility value than a healthy individual in the Swedish HRQL scores used in QALE
 - In the Aldara case a healthy individual has the utility value of 1 and the loss of health if diseased by AK is only 0.007, and as a consequence this patient group is reported to be more healthy than the general population.
 4. $QALE_{pop}$ is based on the general population, while the QALY values in the analysis is based on subgroups
 - Demographic differences between populations may affect both the HRQL and life expectancy
 - Some conditions are gender specific (e.g. prostate cancer for men). Females have a higher life expectancy, but males have a slightly higher HRQL in the data used in this analysis. Since $QALE_{pop}$ is gender-neutral this could cause under- or overestimation of the level of severity.
 - The case of Prolia is affected by this effect and addition to 1. and 3., resulting in a lower value for $QALE_{pop}$ than for the patient population. $QALE_{pat}$ are only females and the utility values are from a different Swedish study than the one used in this thesis.

The variation explained in the four indents are common themes in the data material used in the analysis. To get more reliable severity measurements harmonization of data has to be addressed to the developers of pharmacoeconomic models.

6.2.2 Applications focusing on particular age groups or stages of a condition

Sometimes the pharmaceutical companies only apply for reimbursement for certain age groups or NoMA restricts the indication of the product for only some patient groups. The reason is often that treatment regimens consist of different drugs. Another reason is that there sometimes is only documented effectiveness and/or cost-effectiveness for parts of the population (subpopulations). There is common that diseases have different stages and that drugs are indicated to only be used in parts of a treatment. The models are therefore designed to include only parts of the patient population. The severity ranking of COPD (Onbrez) is a significant case; it is ranked as number 12 in PS and 16 in AS of the 20 cases included in the analysis. The applicant has limited the treatment period for only the first stage of the disease which affects the severity measurement. The result of a severity measurement of that kind of data (not having a lifetime horizon) could mislead to a under- or overestimation of a disease severity if not interpreted carefully. These kind of tailored applications complicate the severity measurements since it leaves out decisive information about the overall severity of a condition. Since the nature of AS includes the time aspect it fails to precisely determine the severity. PS is not that strongly affected, but still fails to give a total impression of the condition due to the data material. This kind of focusing is often done in reimbursement applications, and for both the applicant and NoMA, it is of importance to be specific in the measurement of utilities and time horizon to get valid CUAs and measurements of severity. In relation to the curve in Figure 3 (p. 22), this makes it more challenging to understand the consequences and the severity of a disease. A study conducted by Spencer et al. (2005) indicates that the QALY value for COPD patients given placebo with a median age 63.5 (SD 8.48) was 4.08 ($d = 0.05$) which results in the following severity measures:

- Absolute shortfall: 5.84, would have been ranked as number 9 out of 21
- Proportional shortfall: 59 %, would have been ranked as number 11 out of 21

The discrepancy between these findings and the case of Onbrez supports an argument of including lifetime perspective in evaluations of severity. This could be a request highlighted from NoMA to applicants when developing economic analyses and dossiers for reimbursement assessment.

6.2.3 Patient populations with a high standard deviation in age

In conditions where the patient group has a high standard deviation in median age it could be misleading to use the two methods discussed in this thesis. The case of CML (Glivec) is an example where the patient populations median age is 51 years ranging from 18-70 years old with 22 % of the patients being older than 60. Since the median age of the patient group is of importance in the calculation of AS and PS, a high standard deviation would cause uncertainty in the analysis. The population used in the application is most likely representative for the Norwegian context, but a change in the median age would have noticeable effects on the severity measurement. What makes this case especially interesting is the effectiveness of the drug applied for. Glivec is the only (close to) revolutionary innovation in the study sample (see Table 16 in appendix) with a QALY gain close to five compared to established treatment, according to the applicant. This might get attention from patients, politicians, media and other stakeholders; and having good documentation of both cost-effectiveness and severity would be vital to give reasons for a decision made. For conditions that affect people in all age groups it might be an idea to categorize the patient population in age sub-groups to get a more accurate severity determination. This is especially relevant when using AS where loss of life years is a potent element in the calculation.

6.2.4 Discounting of the level of severity

When deriving QALY values from the different cases it was chosen to use the same discount rate as used by the applicant to counteract biases. In AS, conditions will be valued more severe when using a lower discount rate. This leads to a discrimination of applications where using a higher discount rate, since the diminishing effect on $QALE_{pop}$ and $QALE_{pat}$ is lower. As an action to neutralize this discrimination an attempt to modify the discounting of the established treatment with the discount rate required by NoMA (from 2.5 % to 4 %). This proved that there is not straightforward to do this by first doing a reversed discounting from 2.5 % to zero and then discount with the 4 % rate as recommended by NoMA. The reason is that the discounted QALY value is a sum of calculations within the Markov model, and when simply reversing it we end up with a misleading conclusion.

Another issue about discounting is whether it should be done at all when valuing severity. Although discounting costs and effects at different rates can lead to inconsistencies in reasoning and impossible conclusions (Drummond et al., 2005), the opportunity cost

argument for discounting health gains is hard to understand transferred to the issue of level of severity even if it is basically two sides of the same coin. Theoretical explanations of time preference and discounting is based upon diminishing marginal utility of income, risk aversion and shortsightedness (Richardson, 1999). There is no conformity about discounting of health benefits (Claxton et al., 2011, Gravelle et al., 2007). Nord (2011) recently published a paper where he states that the sensible argument for equal discounting of costs and benefits rooted in microeconomic theory, is not supported by empirical observations of individual's time preferences for health. It is difficult to understand that people will devalue future severity with the same perception as they devalue future consumption when having in mind the characteristics of severity defined by the Lønning Commission: loss of functionality, pain, and discomfort (NOU, 1997). A reasonable and pragmatic approach could be to use only crude severity values (i.e. no discounting) of either QALE or loss of QALE, in contexts where severity is not incorporated in trade-off models. In other circumstances where cost and effects are discounted it sounds more reasonable to also discount severity to counter inconsistencies in reasoning.

6.3 Implications of the severity measures

6.3.1 The valuation of severity by the public

Is it right to bypass the measurement of individual utility and interpret QALYs as measures of social valuation in themselves? When the QALY valuations are generated (e.g. by EQ-5D with TTO) the individuals are asked to value health states for themselves. However, when these data are used in decision making with a societal perspective a discrepancy occurs since the respondents are asked for a personal valuation, not imagining a resource allocation context (Nord, 1994). A person trade-off (PTO) technique would be more precise to measure a societal value of a condition. PTO consists in asking people how many outcomes of one kind they consider equivalent in social value to X outcomes of another kind, e.g. the number of chronically ill people cured which is equivalent to saving 10 lives (Nord, 1995, Pinto Prades, 1997). In reality, there are lots of methods to establish utility values for CUA, and even for EQ-5D values, which is reflected in the included cases. It is necessary to be pragmatic in the world of applied health economics, and to be able to gather adequate HRQL data at all, it seems like it is inevitable to use individual valuation also as a societal preference. Attempts has though been done to transform individual utility values to societal, see Nord (1999). This issue is relevant in the severity discussion because the derived QALY valuations are used in

both of the severity measures in the analysis, which area of application are at a population level. Another issue regarding quantifying of severity is the quantitative relationship between the conditions. Assume that both AS and PS are on a ratio scale. Then we can incorporate that they have properties such as having equidistant points between each of the scale elements, which means that for example multiple myeloma (PS=61) is twice as severe as COPD (PS=28). The same principle would be valid for AS (e.g. AS of 10 is twice as severe as 5). In theory, this sounds like a pragmatic approach, but it probably challenges how the public understands the severity of different conditions, and it also challenges whether or not it is a justification for this interpretation.

6.3.2 Preventive treatments

An interesting finding in the inclusion process was the very low level of measured severity of conditions where a vaccine was the treatment evaluated. None of these was included because of shortcomings of the undertaken analysis. Primary preventive actions, such as mass-vaccination against HPV (all 12 year old girls), are usually offered to big populations, but because few girls actually develop cervical cancer, the QALY gain per vaccinated individual is quite small. This being the case, the QALY gain for the girls without treatment who would have otherwise been diseased has to be divided by all the individuals in the cohort. Both AS and PS are measuring the condition of the risk group or treated individuals and not the condition of the diseased. The result might be understood as counterintuitive since most people would characterize cervical cancer as a severe condition. Hence, this will in itself mean that primary preventive actions will be downgraded if AS or PS are used as measures of severity level. Secondary prevention (methods to diagnose and treat extant disease in early stages before it causes significant morbidity (NLM, 2011)) such as prophylaxis for cardiovascular disease (Brilique) and osteoporosis (Prolia) seems to be characterized by the same pattern, though not so extreme, since these patients already has an indication of increased risk of disease. As described on page 38, the Prolia-case is affected by more than one factor creating biases on the severity of the condition. When the condition is treated with a secondary preventive action, the result is a negative severity value. The difference is illustrated in Figure 7 where the conditions severity is measured at state 1) for primary preventive actions, and in state 2) for secondary preventive actions.



Figure 7 Simplified model of the relationship between disease states

The point of departure in time or disease states is therefore one of the key factors when determining if a condition is characterized as severe or not. All the included cases in this analysis have state 2) or 3) as starting point in their respective pharmacoeconomic models, while in the case of mass-HPV-vaccination of 12-year-old girls; the target population will be in state 1.

Osteoporosis is represented with two cases in the analysis (Prolia and Aclasta) and the main purpose of both is to prevent fractures and could therefore be categorized as a secondary preventive action, since the patient population is indicated as a risk group. Rank wise AS values the conditions slightly more severe than the PS. This is in accordance with the severity definition by the Lønning Commission where prevention is not emphasized (NOU, 1997). However, many believe that prevention is the key concept that can reduce medical costs at the same time that it improves health. Though, a systematic review of the cost-effectiveness literature performed by Cohen et al. (2008) indicates that 80 percent of preventive options add more medical costs than they save. The conditions that are treated with a secondary preventive action have a low severity score in this analysis and may verify the findings in the systematic review.

6.3.3 Historically severe conditions

The comparator in the reimbursement applications are the established treatment which results in a situation where conditions that historically were perceived as very severe, with today's treatment are considered as less severe when undertaking a severity measurement with AS and PS. How severe a condition is would naturally change over time, since new and better treatments are displacing old and less efficient treatments. Consequently we get a dynamic measure of a condition's severity, and ideally AS and PS would be close to zero over time. A new treatment with 100 % effectiveness would for instance lead to zero in severity value for a condition. The treatment for HIV/AIDS is an example of a condition that has had extraordinary improvements the last decades, and a 20-year-old HIV-positive person starting antiretroviral therapy today can expect to live, on average, to the age of 69 (ATCC, 2008). Cardiovascular diseases are a class of diseases that have had remarkable improvements the last 40 years (Figure 8).

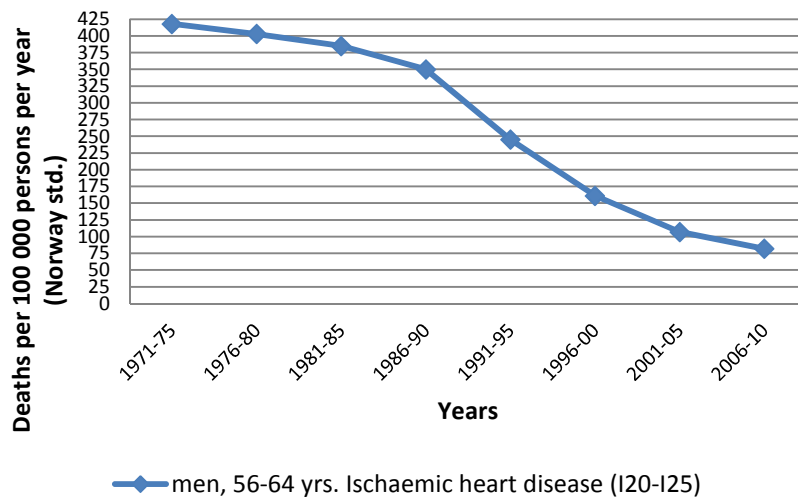


Figure 8 Trend in mortality from ischemic heart disease for Norwegian men in the age groups 45-64 years (FHI, 2012)

This trend is confirmed by the severity of the two applications concerning cardiovascular diseases:

- Brilique: AS = 0.77, PS = 7 %
- Efient: AS = 1.92, PS = 16 %

Compared to the severity values of the other 18 conditions these two are ranked at the lower part of table. This might lead to an imbalance between what the public perceive as severe, and what is considered as severe in health economic evaluations. If financial or other incentives being linked to severity levels, it could have an impact on research and pharmaceutical innovations. The pharmaceutical industry, has expressed skepticism about this way of operationalizing severity. They claim that it would result in an unfavorable shift of where research and development attention is given, and it does not take into account that most medical innovations often are characterized by many small improvements, not revolutions. It might result in that developers end their efforts for making treatments when the profits are “good enough” and not focusing on wiping out diseases completely (GlaxoSmithKline, 2011). Hence, being focused on “get one’s money worth” could lead to discrepancy between health care decision makers and Joe Public. On the other hand, distortion emerges if using a static historical severity measure. Thus, patient groups that are diseased by a historical severe condition would be characterized with an unrealistic level of severity independent of improvements in mortality and quality of life. The 500 % reduction of mortality caused by ischemic heart disease the last 40 years would for example not be given any influence on the severity level. By using a dynamic severity measure, the areas with the greatest potential for

improvements will be emphasized and may give priority. The argument is though in conflict with the wording in §2 in Regulations on the prioritization of health services and the right to health care which states that it is the shortfall of the condition untreated that are given priority (MoH, 2001). The rationale of transferring that legal statement to this topic is though uncertain and most people would probably characterize the consequences as illogical.

6.3.4 Implications of the severity measure procedures

In chapter 4.4 the main features of the principles were described, and this section will do comparisons of how implementable AS and PS are concerning their outcomes. The framework in chapter 2 and 3 will also be used to include aspects of the theories and the policy. Most of the attention will be given to the factor that distinguish them the most: the dimension of time.

AS and PS have in common that they do not have a retrospective perspective measuring experienced health. Other features they share are that they do not state where on the 0-1 scale the condition is located, and there is not possible to separate the two components of QALY and QALE (time and HRQL) to evaluate the ratio. To have specific information about these aspects are valuable for decision makers to get a good understanding of the level of severity, and are an argument for supplementing case papers with additional information. AS and PS do though include experienced life years and expected time of death in two fundamentally different ways:

- AS stress the value of emphasizing young age and the number of lost life years and reduced HRQL
- PS is not that influenced by young age, but is more focused on how near in time death will occur

It was expected before running the analysis that AS would be sensitive to variation in the age of the patient group. When operationalizing the procedure it became clear that the effect of discounting would reduce the age discriminating mechanism. Table 9 and Figure 9 illustrate a condition that cause close to immediate death being measured for different age groups. The max point of the AS scale shrinks from 67 to 20 and narrows the scale. In the undiscounted AS the severity level of a newborn (age zero) is more than eleven times higher than for 80 year olds. For the discounted version which is the used in the analysis, the same relationship is less than four times as severe.

Death at age of	0	10	20	30	40	50	60	70	80	90	100
AS $d = 0$	67.17	58.42	49.64	41.37	33.12	25.33	18.12	11.52	6.07	2.72	1.37
AS $d = 0.04$	21.86	21.01	19.81	18.56	16.76	14.61	11.96	8.64	5.11	2.50	1.33

Table 9 The effect of discounting on the AS scale

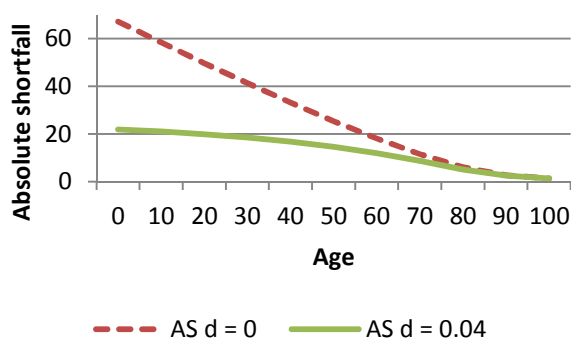


Figure 9 Variation in the range of AS. A 4 % discount rate was used in the analysis

Discounting is a recurring issue which indicates that discounting is an important component of severity measurements, and will be crucial to have clarified this topic with applicants when new reimbursement processes are initiated to be able to do valid severity measurements.

PS has a horizontal age-weighting, i.e. equal weight to all years. If running the same example as done in Table 9 for PS the result would be close to 100 % for every age group independent of discounting rate. Since the outcome of PS is a proportional number this is not surprising, but PS does not completely eliminate the age dimension. In the following table there are three age groups, 70, 80 and 90 where all patients have one QALE left, and are compared to the corresponding general population. This shows how the age aspect influences PS and indicates a much milder inequality in the valuation than AS.

Age	QALE _{pat}	QALE _{pop}	AS	PS
70	1	8.65	7.65	88 %
80	1	5.11	4.11	80 %
90	1	2.50	1.50	60 %

Table 10 How age influences the PS principle

How the outcomes of AS and PS are in accordance with relevant policy and the public's perception of severity level is a complex question. An empirical study conducted by Stolk et al. (2005) investigating the distributional preferences of people, indicates that the equity concept of fair innings correlates most with people's equity interpretations ($r=0.95$).

Spearman correlation was used to analyze the rank order, and a weaker correlation were found for proportional shortfall ($r=0.82$) and severity of illness ($r=-0.65$). This was a Dutch population, but studies have previously shown a high level of accordance between the Dutch and Norwegian population when valuing health (Nord, 1991). As described in section 3.2.2, fair innings is concerned with lifetime health, and both AS and PS have a prospective approach. The difference is that years lived will affect AS valuations, but it does not say anything about the quality of those years. Robberstad (2005) has illustrated the age-weights of the fair innings principle as shown in Figure 10 and called it egalitarian ageism, where people are given gradually less weight in the distribution of scarce health care resources.

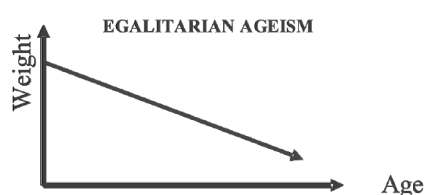


Figure 10 Fair innings age-weighting

This shape has a common feature with Figure 9 and the results from the Netherlands indirectly points at AS being more in line with the expectation of the public what severity issues matter. This could be an argument of making use of AS in severity measurements.

On the other hand, an interpretation of the Lønning Commissions report (NOU, 1997) regards that PS are more in line with the understanding of severity: death risk or loss of physical or psychological functionality, pain and discomfort. No clear indications of age or time considerations. The commission also talks about loss of prognosis and to see the level of severity in relation to the treatment opportunities which implicit indicates that there should be a dimension of time in the priority evaluation. It may be that the minor influence of age in PS is more in line with the idea of severity recommended by the commission.

It is not possible at this point to conclude whether AS or PS is the superior measure of severity level. They both have their characteristics resulting in measuring two different aspects of severity. AS has a broader approach to the topic including factors of equity like egalitarianism. This does not mean that it is an inappropriate severity measure, since this perception of severity has considerable support among people. PS may be characterized as being a more pure severity measure since the influence of the age dimension is reduced. However, the strong element of the prospective health approach makes it vulnerable when a patient group has a high median age. When seeing AS and PS in a decision-making situation

it is clear that they can have a positive impact and contribute to achieve a more comprehensive understanding of a condition in reimbursement applications. While waiting for a clarification from the policymakers this could be an argument of utilizing them in combination, given that the procedures are thoroughly prepared and the users understand their fundamentals. The next section will present an example of how implementation of the severity measures may be done, and what implications that follows.

6.4 Trade-offs

As decision makers want to balance the different priority criteria of health care it is necessary to do trade-offs in a manner that considers both a transparent and consistent process. In the context of this analysis, it is possible to approach the challenge with doing a trade-off between severity and cost-effectiveness. This is exemplified in the two following tables with a quantification of cost-effectiveness as an incremental cost-effectiveness ratio (ICER). The ICER is a measure that reports the additional benefits to be gained from a new therapy (tx) and at what cost (Drummond et al., 2005):

$$(12) \quad ICER = \frac{\text{Difference in costs between } tx_1 \text{ and } tx_2}{\text{Difference in health effects (QALYs) between } tx_1 \text{ and } tx_2}$$

The establishment of severity classes (mild to very severe) is an arbitrary choice made by me to make an impression of the trade-off. The ICER values are calculated by the pharmaceutical companies and are not necessary endorsed by NoMA.

Incremental cost-effectiveness ratio (ICER) (NOK/QALY)						
Severity (Absolute shortfall)		Cost-saving <0	Very good 0-50.000	Good 50 000-250 000	Fairly 250.000-500.000	Poorly > 500.000
	Mild AS: < 1	Mezavant Prolia	Brilique	Onbrez		
	Moderate AS: 1-5	Aclasta	Efient Multaq	Toctino	Ebixa Thalidomide Pharmion	
	Severe AS: 5-10	Firmagon		Revlimid Victoza	Relistor Glivec (CML) Glivec (GIST)	Nexavar
	Very severe AS: > 10		Iressa	Targiniq		

Table 11 Trade-off between cost-effectiveness (ICER) and the level of severity (AS). The threshold value of 500 000 is a hypothetical value derived from a report issued by the Norwegian Knowledge Centre for the Health Services (Wisløff T, 2008). The table is an adapted version of an example used by the Norwegian Directorate of Health in a discussion paper (HDIR, 2011). The Aldara case is excluded from the table because of insufficient cost-effectiveness estimates

Incremental cost-effectiveness ratio (ICER) in NOK			
Product name and condition	ICER	Product name and condition	ICER
Aclasta - Osteoporosis	-439 886	Multaq - Atrial fibrillation	49 604
Aldara - Actinic keratosis	N/A	Nexavar - Hepatocellular carcinoma	679 515
Brilique - Acute coronary syndromes	18 257	Onbrez - Chronic obstructive pulmonary disease (COPD)	154 000
Ebixa - Alzheimer dementia	315 000	Prolia - Osteoporosis	-84 000
Efient - Acute Coronary Syndromes	23 482	Relistor - Opioid-induced constipation in cancer treatment	465 000
Firmagon - Hormone-dependent metastatic prostate cancer	-867 237	Revlimid - Relapse refractory multiple myeloma	247 078
Glivec (CML) - Chronic myelogenous leukemia	384 114	Targiniq - Severe pain opioid induced constipation (P)	23 482
Glivec (GIST) - Gastrointestinal Stromal Tumors	268 400	Thalidomide Pharmion - Multiple myeloma	256 547
Iressa - Non-small-cell lung cancer	35 694	Toctino - Severe chronic hand eczema	136 767
Mezavant - Ulcerative colitis	-287 051	Victoza - Diabetes type 2 obese patients	249 653

Table 12 Incremental cost-effectiveness ratio (ICER) presented in Norwegian Kroner (NOK)

Incremental cost-effectiveness ratio (ICER) (NOK/QALY)						
Severity (Proportional shortfall)		Cost-saving <0	Very good 0-50.000	Good 50 000-250 000	Fairly 250.000-500.000	Poorly > 500.000
	Mild PS: < 25%	Mezavant Aclasta-1 Prolia	Brilique Efient-1 Multaq-1	Toctino-1		
	Moderate PS: 25-50%			Victoza-1 Onbrez+1		
	Severe PS: 50-75%				Glivec (CML) Thalidomide Pharmion+1	
	Very severe PS: 75-100%	Firmagon+1	Iressa	Targiniq Revlimid+1	Glivec (GIST)+1 Relistor+1 Ebixa+2	Nexavar

Table 13 Trade-off between cost-effectiveness (ICER) and the level of severity (PS). The transition of severity class from the AS measurement is marked with + (more severe) and – (less severe)

Hepatocellular cancer (Nexavar) is characterized by being severe/very severe and having a poor cost-effectiveness. The condition of the patient group brings out the characteristics of the two measures; short remaining life expectancy for PS and relative high age for AS. The ICER in the case varies between 550-680' depending on the modeling, and it is expressed by NoMA being above the limit of what is accepted and that the results are highly uncertain. The application was therefore rejected by NoMA. Some may argue that this violates the understanding of severity expressed by the Lønning Commission and the philosophy of the principles of prospective health and severity of illness. What if the analysis of the treatment efficacy was more certain; should the Willingness to Pay (WTP) be higher for this patient group at the expense of other less severe conditions? Nexavar is a life extending cancer drug and recent signals from the Director General of Health, Bjørn Inge Larsen, remarks that the health care authorities need to make some tough choices between the opportunity of making use of expensive life extensive drugs or accommodate for a dignified end of life, making use of scarce health care resources for other groups (Larsen, 2012). Both measures indicate that this is a very severe condition and when taking cost-effectiveness into consideration it seems like NoMA are in line with the suggested trade-off in the tables above.

The condition of Alzheimer dementia (moderate to severe) which is the indication of Ebixa is the condition that has the greatest re-ordering when comparing the AS and PS measures, with a jump from moderate (AS) to very severe (PS). The condition is recognized as a disease that

occurs among elderly and the reimbursement application is focused on patients >80 years old. This change between severity categories is therefore an example of the nature of the fundamental difference between the two measures. In the example made in this thesis, the willingness to pay is therefore remarkably higher when using PS and the opportunities for costly health care technology (drugs) for this patient group will be seriously affected by the approach being used. The danger of ignoring the difference between the measures is illustrated in this case, and is an argument for decision makers to think through what severity approach being used in resource distribution.

Being defined as mild (PS) and moderate severe (AS) with a “good” cost-effectiveness (137 000 per QALY) Tocrino is one of the two included drugs that has been rejected for general reimbursement by NoMA due to uncertain efficiency data and low cost-effectiveness. The certainty of the effectiveness data was criticized by NoMA and might be overestimated by the applicant. If we assume that the drug has an even weaker cost-effectiveness, this can be an indication of a scaled WTP in NoMA's practice where less severe conditions have a lower WTP. It may also be likely that the QALY gain (0.284) achieved by Tocrino (vs. the comparator) would disappear if a stricter valuation was undertaken and the ICER would qualify for rejection.

6.4.1 Double counting

In the trade-off tables at page 50 and 51 we can witness that some conditions like Chronic myelogenous leukemia (CML) might be exposed for a type of double counting in a decision making process. The drug, Glivec, gains as much as 5 QALYs (see appendix iii) and would be eligible for a high WTP. In the study used in the reimbursement application the condition strikes relatively young people with a median age of 51. This qualifies for “severe” in the mentioned tables for both measures. Both the ICER and the severity measures are concerned with a lifetime perspective and will contribute to ageism:

1. The gained life years of the intervention will contribute to a high QALY improvement.
2. The severity measures are concerned with the established treatment and the morbidity and mortality of young individuals result in a considerable loss of QALE.
 - a. especially presented in AS

This means that in cases with an existing high severity because of absence of effective interventions, where the median age is low, and a new effective life-prolonging treatment is introduced; the age factor will have an overlapping effect. Many people would probably

characterize this as reasonable, but in a resource allocation process this can contribute to double counting resulting in a discrimination of other patient groups (e.g. elderly in severe pain with a lower potential for health improvements).

6.5 Limitations

When having an explorative approach, it is challenging to pinpoint variables, data sources or shortcomings of the methodology, since there are a small number of reference points available. Using QALE as done in this thesis has not been identified done before, and the method needs refinement to be successful in a priority situation. A specific issue is to make a model that has a weighted gender composition that matches the model developed by the applicant. Conditions that have homogeneous gender population should therefore be compared to a corresponding population since life expectancy differs between the genders, causing measurement errors.

The data set consists of values derived from both NoMAs own reports and material submitted by pharmaceutical companies (see appendix iii). The values included in the analysis are not necessarily endorsed by NoMA (e.g. ICER values), and should be regarded thereafter.

6.5.1 EQ-5D values used in the QALE procedure

The EQ-5D values used in the QALE-measure may have been affected by several of reasons, e.g. how well the data from the survey represent Stockholm County's adult population. Individuals with poor health may have been underrepresented in the postal questionnaire. Although the non-response rate was moderate (37 %), the authors claim that the results are probably overestimated since it is more problematic for people with a low HRQL to respond (Burström and Rehnberg, 2006). There is no separation of the genders in the QALE model used in this thesis. Since the males in the Swedish survey have throughout a higher HRQL score than the females, we have two factors contributing to the data being too optimistic on behalf of the population. One factor that may stabilize this bias is the higher response-rate by woman and when this HRQL data is matched with the gender neutral life tables it will counteract (maybe too much), since life expectancy for women is significantly higher.

It is documented that EQ-5D is insensitive to some conditions and it may be caused by the relative few, 243, health states. Other similar instruments have defined more than millions of health states and could be more appropriate when measuring severity level. Another disadvantage of EQ-5D is the so-called "ceiling effect". This effect is expressed because large

proportions of respondents indicate no problem and yet many of these respondents report problems on other instruments. It seems that this also result in milder problems getting higher mean values when using EQ-5D. The PS-histogram may be an indicator of this effect. These differences are likely to have an important effect on cost per QALY ratios but the phenomenon has not been thoroughly examined (Brazier and Ratcliffe, 2007). In the context of measuring severity, this may give undesirable effects causing incorrect ratings. Eight out of the 20 cases in the analysis have a severity value that might be characterized as moderate (PS: <30 % AS: <2.2) and according to the criticism above it is reasonable to believe that this is a category of conditions where inaccuracy is present.

7 CONCLUSION

The purpose of this thesis has been to examine how severity measures can be operationalized in applications to the Norwegian Drug Reimbursement Scheme, and what the implications of the alternative measures are. The investigation was carried out with an explorative approach with an aim to give an empirical contribution to the debate concerning how a condition's severity is approached.

A contextual framework was presented including both institutional and theoretical dimensions. There is no existing policy on how severity should be determined in Norway, and governmental agencies such as the Norwegian Medicines Agency, are, therefore, using an arbitrary approach. Health and disease have been approached by a bio statistical theory where health is defined as normal functioning, based on the average health state of the population the individual belongs to (Boorse, 1977). Severity is understood as the magnitude of lost health.

By doing a systematic search to identify severity measures, two QALY-oriented measures were included in the analysis. One is characterized as an absolute measure (AS) that considers both the age and time dimension in addition to physical and psychological discomfort, and mortality. The other is a proportional approach (PS) that assumes that measurement of severity in health should concentrate on the fraction of QALYs that people lose relative to their life expectancy (van de Wetering et al., 2011).

Throughout the thesis, aspects concerning operationalization of the methods were addressed. To achieve valid and reliable measurements there has to be a communicated expectation from NoMA to the applicants about some central issues:

- Harmonization of how utility values are derived
- Use of lifetime horizon in the pharmacoeconomic model
- Handling of subgroups and groups with a high standard deviation in age
- Consistent discounting practices
- Counteraction of double counting between the severity measurement and the CUA

20 different drug reimbursement applications were assessed. A Spearman rank correlation indicated a strong, positive correlation between the two severity measures which was statistically significant at a 0.01 level ($r_s(18) = .933$, $P = .000$). The measures have common features, such as putting a low level of severity on historical severe conditions and conditions

where primary and secondary preventive actions are prescribed. However, there was a significant variation in how the different conditions' severity level was rated. Patient groups with a low median age were emphasized by AS and the opposite mechanism occurred for PS. The findings indicate that the measures are measuring two different aspects of severity, where AS is more concerned by aspects people would characterize as being thoughts of equity. This does not mean that the measure is inferior. Emphasizing lost life years at young age and long-lasting disease are factors that people perceive as severe, and a Dutch study indicate that a close related approach to AS (fair innings), are more in line with peoples preferences than PS (Stolk et al., 2005). The findings in this study therefore indicate that both severity measures may improve consistency in decision making, in addition to achieve a more comprehensive understanding of a condition. The methodology used in this thesis has not been seen used for the same purpose, and it would be interesting to see more research done on applying severity measures to real-life data in order to develop new and refine existing principles.

REFERENCES

- ALBERTSEN, P. C., HANLEY, J. A., GLEASON, D. F. & BARRY, M. J. 1998. Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer. *JAMA*, 280, 975-80.
- ANDERSON, J. L., ADAMS, C. D., ANTMAN, E. M., BRIDGES, C. R., CALIFF, R. M., CASEY, D. E., JR., CHAVEY, W. E., 2ND, FESMIRE, F. M., HOCHMAN, J. S., LEVIN, T. N., LINCOFF, A. M., PETERSON, E. D., THEROUX, P., WENGER, N. K., WRIGHT, R. S., SMITH, S. C., JR., JACOBS, A. K., HALPERIN, J. L., HUNT, S. A., KRUMHOLZ, H. M., KUSHNER, F. G., LYTLE, B. W., NISHIMURA, R., ORNATO, J. P., PAGE, R. L. & RIEGEL, B. 2007. ACC/AHA 2007 guidelines for the management of patients with unstable angina/non-ST-Elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 2002 Guidelines for the Management of Patients With Unstable Angina/Non-ST-Elevation Myocardial Infarction) developed in collaboration with the American College of Emergency Physicians, the Society for Cardiovascular Angiography and Interventions, and the Society of Thoracic Surgeons endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation and the Society for Academic Emergency Medicine. *J Am Coll Cardiol*, 50, e1-e157.
- ATCC, T. A. T. C. C. 2008. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. *Lancet*, 372, 293-9.
- BASSAND, J. P., HAMM, C. W., ARDISSINO, D., BOERSMA, E., BUDAJ, A., FERNANDEZ-AVILES, F., FOX, K. A., HASDAI, D., OHMAN, E. M., WALLENTIN, L. & WIJNS, W. 2007. Guidelines for the diagnosis and treatment of non-ST-segment elevation acute coronary syndromes. *Eur Heart J*, 28, 1598-660.
- BOORSE, C. 1977. Health as a theoretical concept. *Philosophy of Science*, 44, 542-573.
- BRANDEN, E. 2008. [Lung cancer in women. A growing diagnostic and therapeutic challenge]. *Lakartidningen*, 105, 1653-5.
- BRAZIER, J. & RATCLIFFE, J. 2007. *Measuring and valuing health benefits for economic evaluation*, Oxford University Press.
- BRAZIER, J. E., YANG, Y., TSUCHIYA, A. & ROWEN, D. L. 2010. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ*, 11, 215-25.
- BURSTRÖM, K. & REHNBERG, C. 2006. *Hälsorelaterad livskvalitet i Stockholms län 2002* [Online]. Available: [http://www.folkhalsoguiden.se/upload/J%c3%a4mlik%20h%c3%a4lsa/H%c3%a4lsorelaterad%20livskvalitet%20i%20Stockholms%20l%c3%a4n%202002%20\(2006_1\).pdf](http://www.folkhalsoguiden.se/upload/J%c3%a4mlik%20h%c3%a4lsa/H%c3%a4lsorelaterad%20livskvalitet%20i%20Stockholms%20l%c3%a4n%202002%20(2006_1).pdf).
- CLAXTON, K., PAULDEN, M., GRAVELLE, H., BROUWER, W. & CULYER, A. J. 2011. Discounting and decision making in the economic evaluation of health-care technologies. *Health Econ*, 20, 2-15.
- COHEN, J. T., NEUMANN, P. J. & WEINSTEIN, M. C. 2008. Does preventive care save money? Health economics and the presidential candidates. *N Engl J Med*, 358, 661-3.
- CUADRAS-MORATO, X., PINTO-PRADES, J. L. & ABELLAN-PERPINAN, J. M. 2001. Equity considerations in health care: the relevance of claims. *Health Econ*, 10, 187-205.
- DOLAN, P., SHAW, R., TSUCHIYA, A. & WILLIAMS, A. 2005. QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ*, 14, 197-208.
- DRUMMOND, M. F., SCULPHER, M. J., TORRANCE, G. W., O'BRIEN, B. J. & STODDART, G. L. 2005. *Methods for the economic evaluation of health care programmes*, Oxford, Oxford University Press.
- DURIE, B. G. 2001. The epidemiology of multiple myeloma. *Semin Hematol*, 38, 1-5.

- EGAN, T. M. 2002. QALYs or quackery? The quagmire of quantifying the cost of breathing. *J Thorac Cardiovasc Surg*, 123, 406-8.
- EUROQOL. 2012. *What is EQ-5D* [Online]. Available: <http://www.euroqol.org/eq-5d/what-is-eq-5d.html> [Accessed 05.05.2012 2012].
- FADERL, S., TALPAZ, M., ESTROV, Z. & KANTARJIAN, H. M. 1999. Chronic myelogenous leukemia: biology and therapy. *Ann Intern Med*, 131, 207-19.
- FHI, N. F. 2011. Retningslinjer for legemiddeløkonomiske analyse - høringsuttalelse.
- FHI, N. F. 2012. Causes of death, european short-list – deaths per 100 000 persons per year (Norway std.), Ischaemic heart disease (I20-I25)
http://www.norgeshelsa.no/norgeshelsa/index.jsp?study=http%3A%2F%2F10.0.3.22%3A80%2Fobj%2Fstudy%2Fhjertedod&mode=cube&v=2&context=http%3A%2F%2F10.0.3.22%3A80%2Fobj%2Fcatalog%2Fnorgeshelsa&cube=http%3A%2F%2F10.0.3.22%3A80%2Fobj%2Fcube%2Fhjertedod_C1&top=yes&language=en Nasjonalt Folkehelseinstitutt FHI.
- FOWLER, J. 2008. Chronic hand eczema: a prevalent and challenging skin condition. *Cutis*, 82, 4-8.
- FOX-RUSHBY, J. & CAIRNS, J. 2005. *Economic Evaluation*, Open University Press.
- GAUTHIER, A., KANIS, J. A., MARTIN, M., COMPSTON, J., BORGSTROM, F., COOPER, C. & MCCLOSKEY, E. 2011. Development and validation of a disease model for postmenopausal osteoporosis. *Osteoporos Int*, 22, 771-80.
- GELDMACHER, D. S., FROLICH, L., DOODY, R. S., ERKINJUNTTI, T., VELLAS, B., JONES, R. W., BANERJEE, S., LIN, P. & SANO, M. 2006. Realistic expectations for treatment success in Alzheimer's disease. *J Nutr Health Aging*, 10, 417-29.
- GLAXOSMITHKLINE 2011. Høring av nye retningslinjer for legemiddeløkonomiske analyser. In: GLAXOSMITHKLINE (ed.) *Arbeid med reviderte retningslinjer for legemiddeløkonomiske analyser*.
<http://www.google.no/url?sa=t&rct=j&q=glaxosmithkline%20legemiddelverket%20h%C3%B8ring&source=web&cd=2&ved=0CCgQFjAB&url=http%3A%2F%2Fwww.legemiddelverket.no%2Fupload%2F158848%2FH%25C3%25B8ringsuttalelse%2520fra%2520GSK.pdf&ei=CJVST8qsF-iF4gTxuoHAAG&usg=AFQjCNH1PJAK48RMW3Gwha6FJjmr2EmBLg> Statens Legemiddelverk.
- GO, A. S., HYLEK, E. M., PHILLIPS, K. A., CHANG, Y., HENAULT, L. E., SELBY, J. V. & SINGER, D. E. 2001. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA*, 285, 2370-5.
- GRAVELLE, H., BROUWER, W., NIESSEN, L., POSTMA, M. & RUTTEN, F. 2007. Discounting in economic evaluations: stepping forward towards optimal decision rules. *Health Econ*, 16, 307-17.
- GRAY, A. M., CLARKE, P. M., WOLSTENHOLME, J. & WORDSWORTH, S. 2011. *Applied Methods of Cost-Effectiveness Analysis in Healthcare*, Oxford University Press.
- HAGEN G, W. T., FALCH J, LOFTHUS C, FRIHAGEN F, WEENSAAS K-A, GRANUM L, NEVJAR J, KRISTIANSEN IS, KLEMP M, 2010. Efficacy and costeffectiveness of alendronate for the prevention of fractures in postmenopausal women in Norway. Report from the Norwegian Knowledge Centre for the Health Services nr. 5-2010.
- HAGINO, C., BOSCARIOL, J., DOVER, L., LETENDRE, R. & WICKS, M. 1998. Before/after study to determine the effectiveness of the align-right cylindrical cervical pillow in reducing chronic neck pain severity. *J Manipulative Physiol Ther*, 21, 89-93.
- HANSSON, L. F., NORHEIM, O. F. & RUYTER, K. W. 1994. Equality, explicitness, severity, and rigidity: the Oregon plan evaluated from a Scandinavian perspective. *J Med Philos*, 19, 343-66.
- HDIR, D. O. H. 2011. Samfunnsøkonomiske analyser i helsesektoren - en veileder. In: HELSEDIREKTORATET (ed.) *Helsedirektoratet*.
<http://www.helsedirektoratet.no/Om/hoyringar/Sider/samfunnsokonomiske-analyser-i-helsesektoren.aspx> Helsedirektoratet.
- JIA, H., ZACK, M. M. & THOMPSON, W. W. 2011. State Quality-Adjusted Life Expectancy for U.S. adults from 1993 to 2008. *Qual Life Res*, 20, 853-63.

- JIA, H., ZACK, M. M., THOMPSON, W. W. & DUBE, S. R. 2012. Quality-adjusted life expectancy (QALE) loss due to smoking in the United States. *Qual Life Res.*
- LARSEN, B. I. 2012. De vanskelige valgene. *Aftenposten*.
- LEE, H. Y., HWANG, J. S., JENG, J. S. & WANG, J. D. 2010. Quality-adjusted life expectancy (QALE) and loss of QALE for patients with ischemic stroke and intracerebral hemorrhage: a 13-year follow-up. *Stroke*, 41, 739-44.
- LLOVET, J. M., BURROUGHS, A. & BRUIX, J. 2003. Hepatocellular carcinoma. *Lancet*, 362, 1907-17.
- LUCAS, R., MCMICHAEL, T., SMITH, W., ARMSTRONG, B. 2006. Solar Ultraviolet Radiation: Global burden of disease from solar ultraviolet radiation. *WHO Library Cataloguing-in-Publication Data*.
- LUNDBERG, L., JOHANNESON, M., ISACSON, D. & BORGQUIST, L. 1999. Health-state utilities in a general population in relation to age, gender and socioeconomic factors. *The European Journal of Public Health*, 9, 211-217.
- MCKHANN, G., DRACHMAN, D., FOLSTEIN, M., KATZMAN, R., PRICE, D. & STADLAN, E. M. 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 34, 939-44.
- MOH, M. O. H. A. C. S. 2001. FOR 2000-12-01 nr 1208: Forskrift om prioritering av helsetjenester, rett til nødvendig helsehjelp fra spesialisthelsetjenesten, rett til behandling i utlandet og om klagenemnd (prioriteringsforskriften). In: OMSORGSDEPARTEMENTET, H. H.-O. (ed.).
- MOH, M. O. H. A. C. S. 2012. *Refusjon av legemidler ved langvarig sykdom - blåreseptforskriften* [Online]. Ministry of Health and Care Services. Available: <http://www.regjeringen.no/nb/dep/hod/tema/legemidler/forhandsgodkjent-refusjon-av-legemidler-.html?id=434582>.
- MOHS, R. C. & COHEN, L. 1988. Alzheimer's Disease Assessment Scale (ADAS). *Psychopharmacol Bull*, 24, 627-8.
- MORLAND, B., RINGARD, A. & ROTTINGEN, J. A. 2010. Supporting tough decisions in Norway: a healthcare system approach. *Int J Technol Assess Health Care*, 26, 398-404.
- MOVIK E, R. T., LINNESTAD KK, HOFMANN, B, HARBOE I, KLEMP, M 2009. Metylnaltrekson ved forstoppelse i kreftbehandling. *Rapport fra Kunnskapssenteret nr 23-2009*.
- MUENNIG, P., LUBETKIN, E., JIA, H. & FRANKS, P. 2006. Gender and the burden of disease attributable to obesity. *Am J Public Health*, 96, 1662-8.
- NHLBI, U. S. D. O. H. H. S. N. 2012. *What Is COPD?* [Online]. National Heart Lung and Blood Institute, U.S. Department of Health & Human Services: U.S. Department of Health & Human Services.
- NICE 2008. Report on NICE Citizens Council meeting Quality Adjusted Life Years (QALYs) and the severity of illness.
- NLM, N. L. O. M. 2011. National Library of Medicine - Medical Subject Headings Secondary Prevention. http://www.nlm.nih.gov/cgi/mesh/2011/MB_cgi?mode=&term=Secondary+Prevention: National Library of Medicine NLM.
- NOKC, T. N. K. C. F. T. H. S. 2011. *Slik oppsummerer vi forskning*, <http://www.kunnskapssenteret.no/Verkt%C3%B8y/Slik+oppsummerer+vi+forskning.2139.cms>, The Norwegian Knowledge Centre for the Health Services (NOKC).
- NOMA. 2005. *Norwegian guidelines for pharmacoeconomic analysis in connection with applications for reimbursement* [Online]. Available: http://www.legemiddelverket.no/templates/InterPage_25644.aspx?filterBy=CopyToGeneral.
- NOMA. 2011a. *Blåreseptnemndas oppgaver* [Online]. Available: http://www.legemiddelverket.no/templates/InterPage_79343.aspx?filterBy=CopyToGeneral.
- NOMA. 2011b. *Retningslinjer for legemiddeløkonomiske analyser på høring* [Online]. http://legemiddelverket.no/templates/InterPage_83076.aspx?filterBy=CopyToPharma:

- NOMA. Available:
http://legemiddelverket.no/templates/InterPage_83076.aspx?filterBy=CopyToPharma
 2012].
- NOMA 2012a. Multaq på blå resept.
http://www.legemiddelverket.no/templates/InterPage_83464.aspx.
- NOMA 2012b. Reviderte retningslinjer for legemiddeløkonomiske analyser. In: NOMA (ed.).
http://www.legemiddelverket.no/templates/InterPage_83495.aspx?filterBy=:NOMA.
- NORD, E. 1991. EuroQol: health-related quality of life measurement. Valuations of health states by the general public in Norway. *Health Policy*, 18, 25-36.
- NORD, E. 1993a. [Health care politicians are not concerned about maximum health gain per crown]. *Tidsskr Nor Laegeforen*, 113, 1371-3.
- NORD, E. 1993b. The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Health Policy*, 24, 227-38.
- NORD, E. 1994. The QALY--a measure of social value rather than individual utility? *Health Econ*, 3, 89-93.
- NORD, E. 1995. The person-trade-off approach to valuing health care programs. *Med Decis Making*, 15, 201-8.
- NORD, E. 1999. *Cost-value analysis in health care: making sense out of QALYs*, Cambridge University Press.
- NORD, E. 2005. Concerns for the worse off: fair innings versus severity. *Soc Sci Med*, 60, 257-63.
- NORD, E. 2011. Discounting future health benefits: the poverty of consistency arguments. *Health Econ*, 20, 16-26.
- NOU 1997. Prioritering på ny. In: AFFAIRS, M. O. S. A. H. (ed.).
- OLSEN, J. A. 1997. Theories of justice and their implications for priority setting in health care. *J Health Econ*, 16, 625-39.
- OLSEN, J. A. 2006. *Helseøkonomi: effektivitet og rettferdighet*, Cappelen akademisk forl.
- OLSEN, J. A., RICHARDSON, J., DOLAN, P. & MENZEL, P. 2003. The moral relevance of personal characteristics in setting health care priorities. *Soc Sci Med*, 57, 1163-72.
- PAI, R. G. & VARADARAJAN, P. 2007. Prognostic significance of atrial fibrillation is a function of left ventricular ejection fraction. *Clin Cardiol*, 30, 349-54.
- PETROU, S. & RENTON, A. 1993. The QALY: a guide for the public health physician. *Public Health*, 107, 327-36.
- PINTO PRADES, J. L. 1997. Is the person trade-off a valid method for allocating health care resources? *Health Econ*, 6, 71-81.
- PREEDY, V. R. & WATSON, R. R. 2010. *Handbook of Disease Burdens and Quality of Life Measures*, New York, NY, Springer New York.
- RAJKUMAR, S. V. & KYLE, R. A. 2005. Multiple myeloma: diagnosis and treatment. *Mayo Clin Proc*, 80, 1371-82.
- REVICKI, D. A. & WOOD, M. 1998. Patient-assigned health state utilities for depression-related outcomes: differences by depression severity and antidepressant medications. *J Affect Disord*, 48, 25-36.
- RICHARDSON, J. 1999. Age Weighting and Discounting: What are the Ethical Issues?
<http://www.buseco.monash.edu.au/centres/che/pubs/wp109.pdf>: Centre for Health Program Evaluation Monash University.
- RICHARDSON, J. R., MCKIE, J., PEACOCK, S. J. & IEZZI, A. 2011. Severity as an independent determinant of the social value of a health service. *Eur J Health Econ*, 12, 163-74.
- RIVLIN, M. M. 2000. Why the fair innings argument is not persuasive. *BMC Med Ethics*, 1, E1.
- ROBBERSTAD, B. 2005. QALYs vs DALYs vs LYs gained: What are the differences, and what difference do they make for health care priority setting? *Norsk Epidemiologi*, 15, 183-191.
- ROMANO, P. S. & CHAN, B. K. 2000. Risk-adjusting acute myocardial infarction mortality: are APR-DRGs the right tool? *Health Serv Res*, 34, 1469-89.

- SASSI, F. 2006. Calculating QALYs, comparing QALY and DALY calculations. *Health Policy Plan*, 21, 402-8.
- SHAH, K. K. 2009. Severity of illness and priority setting in healthcare: a review of the literature. *Health Policy*, 93, 77-84.
- SIROHI, B. & POWLES, R. 2004. Multiple myeloma. *Lancet*, 363, 875-87.
- SONNENBERG, F. A. & BECK, J. R. 1993. Markov models in medical decision making: a practical guide. *Med Decis Making*, 13, 322-38.
- SPENCER, M., BRIGGS, A. H., GROSSMAN, R. F. & RANCE, L. 2005. Development of an economic model to assess the cost effectiveness of treatment interventions for chronic obstructive pulmonary disease. *Pharmacoeconomics*, 23, 619-37.
- SSB. 2011. *Population statistics. Deaths. 2010* [Online]. http://www.ssb.no/english/subjects/02/02/10/dode_en/: Statistics Norway. Available: http://www.ssb.no/english/subjects/02/02/10/dode_en/ [Accessed 20.02.2012].
- STOLK, E. A., PICKEE, S. J., AMENT, A. H. & BUSSCHBACH, J. J. 2005. Equity in health care prioritisation: an empirical inquiry into social value. *Health Policy*, 74, 343-55.
- STOLK, E. A., VAN DONSELAAR, G., BROUWER, W. B. & BUSSCHBACH, J. J. 2004. Reconciliation of economic concerns and health policy: illustration of an equity adjustment procedure using proportional shortfall. *Pharmacoeconomics*, 22, 1097-107.
- TCCFA, T. C. S. A. C. F. O. A. 2011. *About Ulcerative Colitis & Proctitis* [Online]. <http://www.ccfa.org/info/about/ucp/>: The Crohn's and Colitis Foundation of America. [Accessed 11.03.2012].
- UBEL, P. A. 1999. How stable are people's preferences for giving priority to severely ill patients? *Social Science & Medicine*, 49, 895-903.
- UHS-PRUITTCORPORATION. 2012. *COPD and Lung Function* [Online]. <http://www.uhs-pruitt.com/Resources.aspx?id=1234>: UHS-Pruitt Corporation. Available: <http://www.uhs-pruitt.com/Resources.aspx?id=1234>.
- VAN DE WETERING, E. J., STOLK, E. A., VAN EXEL, N. J. & BROUWER, W. B. 2011. Balancing equity and efficiency in the Dutch basic benefits package using the principle of proportional shortfall. *Eur J Health Econ*.
- WHO 2003. Prevention and management of osteoporosis. *World Health Organ Tech Rep Ser*, 921, 1-164, back cover.
- WIDIMSKY, P., WIJNS, W., FAJADET, J., DE BELDER, M., KNOT, J., AABERGE, L., ANDRIKOPOULOS, G., BAZ, J. A., BETRIU, A., CLAEYS, M., DANCHIN, N., DJAMBAZOV, S., ERNE, P., HARTIKAINEN, J., HUBER, K., KALA, P., KLINCEVA, M., KRISTENSEN, S. D., LUDMAN, P., FERRE, J. M., MERKELY, B., MILICIC, D., MORAIS, J., NOC, M., OPOLSKI, G., OSTOJIC, M., RADOVANOVIC, D., DE SERVI, S., STENESTRAND, U., STUDENCAN, M., TUBARO, M., VASILJEVIC, Z., WEIDINGER, F., WITKOWSKI, A. & ZEYMER, U. 2010. Reperfusion therapy for ST elevation acute myocardial infarction in Europe: description of the current situation in 30 countries. *Eur Heart J*, 31, 943-57.
- WILD, S., ROGLIC, G., GREEN, A., SICREE, R. & KING, H. 2004. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care*, 27, 1047-53.
- WILLIAMS, A. 1997. Intergenerational equity: an exploration of the 'fair innings' argument. *Health Econ*, 6, 117-32.
- WISLØFF T, N. O., HALVORSEN S, SELMER RM, KRISTIANSEN IS. 2008. Kostnader og leveårsgevinster ved medikamentell primærforebygging av hjertekarsykdom. Rapport Nr 34-2008. <http://www.kunnskapssenteret.no/Publikasjoner/Kostnader+og+leve%C3%A5rsgevinster+ved+medikamentell+prim%C3%A6rforebygging+av+hjertekarsykdom.4786.cms>: Norwegian Knowledge Centre for the Health Services.

APPENDIX

i. Assessed reimbursement applications

Trade name	Case #	Application received	Company	Indication	Type of application	Included/ excluded	Conclusion	Comments
Aclasta	09/06043	28.04.2009	Novartis	Osteoporosis	Indication expansion	Included	\$ 2 with conditions	
Aldara	09/16142	16.11.2009	Meda AS	Actinic keratosis	Indication expansion	Included	\$ 2 with conditions	
Beriner	09/17757	07.05.2010	CSL Behring AB	Hereditary angioedema	New active agent	Excluded	\$2	No severity info
Brilique	11/02350	31.01.2011	AstraZeneca AS	Acute coronary syndromes	New active agent	Included	\$2	
Canoderm 5% k	09/08896	01.07.2009	Aco Hud Nordic AB	atopic dermatitis eximia	New active agent	Excluded	Rejected	Ref: conclusion
Cervarix	10/06086	03.03.2010	GlaxoSmithKline	HPV vaccine	New active agent	Excluded	\$14-32 transferred HOD	Vaccine, no sev
Crestor	09/12940	07.10.2009	AstraZeneca AS	Hypercolosterima	New active agent	Excluded		Using LYG
Ebixa	10/10437	07.05.2010	H. Lundbeck	Alzheimer dementia	New active agent	Included	\$ 2 with conditions	
Eflent	09/03823	05.03.2009	Eli Lilly	Acute Coronary Syndromes	New active agent	Included	\$ 2 with conditions	
EpiPen/EpiPen	09/08921	01.07.2009	ALK Sverige AB	Prophylactic shock, allergics	New active agent	Excluded	\$ 2 with conditions	CMA
Exforge HTC	09/15631	09.11.2009	Novartis Norge AS	Hypertension	New combination	Excluded	\$2	No CUA
Ezetrol	09/03704	03.03.2009	MSD Norge AS	Atherosclerotic	New active agent	Excluded	\$ 2 with conditions	No QALY weight
Firmagon	09/10247	24.07.2009	Ferring	Hormone-dependent metastatic prostate cancer	New active agent	Included	\$ 2 with conditions	
Gardasil	09/16212	18.11.2009	Sanofi Pasteur MSD AS	HPV vaccine	New active agent	Excluded	\$14-32 transferred to HOD	No severity info
Glivec	09/03012	18.02.2009	Novartis	GIST and CML	New active agent	Included	sendt innstilling	
Grazax	09/02619	10.02.2009	ALK Abelló	Rhinitis	Indication expansion	Excluded		CMA
Iressa	09/18176	28.12.2009	AstraZeneca AS	Non-small-cell lung cancer	New active agent	Included	\$14-32 transfer to HOD	
Januvia	09/05261	03.04.2009	MSD	diabetes T2	New active agent	Excluded	\$ 2 with conditions	No QALY values
Klexane	09/00219	07.01.2009	sanofi aventis	Trombosis profylaxis		Excluded		"short" applicat
Levemir	10/09779	16.04.2010	Novo Nordisk	diabetes T2	Indication expansion	Excluded	No to \$ 2	Too simple anal
Mezavant	09/05273	03.04.2009	Shire pharmaceutical contr	Ulcerative colitis	Ny formulering/legemiddelfo	Included		
Multaq	10/11831	31.05.2010	Sanofi-aventis AS	Atrial fibrillation	New active agent	Included		
Nexavar	09/06842	14.05.2009	Bayer AB	Hepatocellular carcinoma	New active agent	Included	Rejected	
Onbrez	10/05025	08.02.2010	Novartis Norge AS	Chronic obstructive pulmonary disease (COPD)	New active agent	Included	\$ 2 kode ICPC R95, ICD J43,J44	
Onglyza	09/15648	03.11.2009	AstraZeneca AS	Diabetes mellitus II	New active agent	Excluded	\$ 2 with conditions	CMA
OsvaRen	10/12603	14.06.2010	Vingmed AS	Dialyse	Ny formulering/legemiddelfo	Excluded	\$ 2 with conditions	CMA
Pradaxa	09/06401	06.05.2009	sanofi aventis	Trombosis profylaxis	New active agent	Excluded	\$ 2 with conditions	CMA
Prolia	10/13347	24.06.2010	GlaxoSmithKline	Osteoporosis	New active agent	Included	\$ 2 with conditions	
Relistor	10/10640	10.05.2010	Pfizer AS	Opioid-induced constipation in cancer treatment	New active agent	Included	\$ 2 with conditions	
Renvela	10/04263	09.02.2010	Genzyme	hyperphosphatemia	Indication expansion	Excluded	\$2, kun refusjon til dialys	No CUA
Revlimid	10/04668	15.02.2010	Celgene Europe Limited	Relapse refractory multiple myeloma	New active agent	Included	\$ 2 with conditions	
Sutent	10/19410	04.10.2010	Pfizer	Metastatic kidney cancer	Indication expansion	Excluded	\$ 2 with conditions	No CUA
Symbicort	09/04074	10.03.2009	AstraZeneca AS	COPD	New active agent	Excluded	Sent recommendation	No CUA
Targiniq	09/17010	02.12.2009	Mundipharma	Severe pain opioid induced constipation	New combination	Included	\$ 2 with conditions	
Tasigna	11/04422	01.03.2011	Novartis	Leukemia	New active agent	Excluded	\$ 2	CMA
Thalidomide Ph	09/04989	24.03.2009	Celgene Europe Limited	Multiple myeloma	New active agent	Included		
Toctino	09/16349	05.07.2010	Basilea Pharmaceuticals	Severe chronic hand eczema	New active agent	Included	Rejected	
Tredaptive	09/09266	07.07.2009	MSD Norge AS	Dyslipidemia	New active agent	Excluded	No to \$ 2	Rejected model
Trobalt	11/06762	19.04.2011	GlaxoSmithKline	Epilepsia	New active agent	Excluded	\$ 2	CMA
Victoza	09/17136	03.12.2009	Novo nordisk Skandinavia	Diabetes type 2 obese patients	New active agent	Included	\$ 2 with conditions	
Votrient	10/14295	12.07.2010	GlaxoSmithKline	Kidney cancer	New active agent	Excluded	Rejected	CMA
Xarelto	09/08239	15.06.2009	Bayer AB	Trombosis profylaxis	New active agent	Excluded	\$ 2 with conditions	Insufficient CUA
Zarzio	10/21928	22.09.2010	Sandoz	Neutropenia	Generika	Excluded	\$ 2	CMA ("kort/lang
				(The Glivec application is split into two				
CUA = Cost utility analysis		Included	19					
CMA = Cost min analysis		Excluded	24					

Table 14 Overview of the assessed applications

ii. EQ-5D utility values

Tabell 3: Procentuell andel av befolkningen som rapporterar inga, måttliga respektive svåra problem per EQ-5D dimension, livskvalitetsvikt EQ-5D index (medelvärde+standardavvikelse), självskattad hälsa EQ VAS (medelvärde+standardavvikelse, medianvärde +kvartiler), män och kvinnor per åldersgrupp, Stockholms län 2002.

EQ-5D dimension	Åldersgrupp (män och kvinnor) (n=30308)												
	18-24 år (n=2384)	25-29 år (n=2497)	30-34 år (n=3071)	35-39 år (n=3160)	40-44 år (n=2758)	45-49 år (n=2539)	50-54 år (n=2794)	55-59 år (n=3221)	60-64 år (n=2384)	65-69 år (n=1735)	70-74 år (n=1483)	75-79 år (n=1275)	80-84 år (n=1007)
Rörlighet													
Inga problem	97,7	97,2	96,9	95,2	93,2	92,2	90,3	87,6	84,0	82,4	75,1	69,4	55,8
Måttliga problem	2,2	2,8	3,0	4,7	6,6	7,7	9,6	12,3	15,8	17,5	24,4	29,9	42,9
Svåra problem	0,1	0,0	0,1	0,1	0,2	0,1	0,1	0,1	0,2	0,1	0,5	0,7	1,3
Hygien													
Inga problem	99,5	99,7	99,5	98,9	98,6	98,5	98,4	98,5	97,6	97,4	95,9	93,7	89,0
Måttliga problem	0,3	0,2	0,4	1,0	1,0	1,3	1,5	1,3	2,0	2,2	3,0	5,0	8,1
Svåra problem	0,2	0,1	0,1	0,1	0,4	0,2	0,1	0,2	0,4	0,4	1,1	1,3	2,9
Huvudsakliga aktiviteter													
Inga problem	94,1	93,9	93,5	91,6	90,2	89,2	88,4	87,5	84,8	89,0	86,2	82,0	71,4
Måttliga problem	5,1	5,3	5,8	7,3	7,9	8,7	9,6	10,5	12,6	10,0	11,4	15,3	22,9
Svåra problem	0,8	0,8	0,7	1,1	1,9	2,1	2,0	2,0	2,6	1,0	2,4	2,7	5,7
Smärtor/besvär													
Inga problem	70,4	68,4	66,0	60,5	57,2	51,5	47,7	44,5	42,4	43,4	38,1	34,3	31,9
Måttliga problem	28,2	30,2	32,3	36,9	38,7	44,2	47,7	50,4	53,2	52,8	57,2	60,9	61,1
Svåra problem	1,4	1,4	1,7	2,6	4,1	4,3	4,6	5,1	4,4	3,8	4,7	4,8	7,0
Oro/nedstämdhet													
Inga problem	54,7	57,1	62,1	61,8	61,1	62,3	61,4	63,5	68,4	70,2	68,4	66,4	61,8
Måttliga problem	40,9	39,6	35,1	35,0	35,2	34,3	35,7	33,7	29,1	28,1	29,9	31,0	36,2
Svåra problem	4,4	3,3	2,8	3,2	3,7	3,4	2,9	2,8	2,5	1,7	1,7	2,6	2,0
Livskvalitetsvikt EQ-5D index													
Medelvärde	0,85	0,86	0,86	0,84	0,83	0,81	0,81	0,80	0,80	0,81	0,78	0,76	0,71
(standardavvikelse)	(0,19)	(0,18)	(0,18)	(0,20)	(0,23)	(0,22)	(0,22)	(0,23)	(0,23)	(0,21)	(0,23)	(0,25)	(0,28)
Självskattad hälsa EQ VAS													
Medelvärde	78,3	78,3	79,0	77,7	76,7	76,9	76,9	75,8	76,3	78,0	74,3	70,3	65,1
(standardavvikelse)	(17,7)	(17,3)	(16,7)	(17,9)	(19,3)	(18,7)	(19,4)	(19,7)	(20,3)	(18,1)	(19,4)	(20,3)	(21,5)
Självskattad hälsa EQ VAS													
Medianvärde	80	80	80	80	80	80	80	80	80	80	80	75	70
(nedre, övre kvartil)	(70, 90)	(70, 90)	(70, 90)	(70, 90)	(70, 90)	(70, 90)	(70, 90)	(69), (90)	(70), (90)	(70), (90)	(65), (90)	(59), (85)	(50), (80)
Antal per åldersgrupp EQ VAS	2041	2216	2721	2805	2444	2298	2519	2937	2167	1537	1276	1078	825

Table 15 Swedish EQ-5D values

iii. Severity calculation

R = Reimbursement report S = Application from company		V = Appendix to application A = Study used in application		X = Assumption, missing data N/A = Not available		AS = Absolute shortfall PS = Proportional shortfall		H = Health economic model (manually run)									
Total applications:		20															
																Discount rate	
																4.0 %	
Product name and condition	QALY established treatment	QALY new treatment	Δ QALY	Method	Discount rate	Population	ICER	Median age / start age in model	Time horizon	QALE	AS	PS	Discounted QALE	AS (4 % disc.)	PS (4 % disc.)	Disc. adjusted AS	Disc. adjusted PS
Iressa	0.870	0.990	0.120	EQ-5D (VAS/SG)	4.0 %	UK	kr 35 694	57	Lifetime (6)	19.752	18.882	96 %	12.338	11.468	93 %	11.47	93 %
Non-small-cell lung cancer	R	R	R		R	S	R	R	S								
Brilique	9.681	9.8138	0.133	EQ-5D	3.0 %	International	kr 18 257	67	Lifetime (40)	12.947	3.266	25 %	9.965			0.28	3 %
Acute coronary syndromes	R	R	R	V	R	S		V	S								
Tocino	11.548	11.832	0.284	EQ-5D (mapping)	2.5 %	Northern Europe	kr 136 767	48	22	17.507	5.959	34 %	13.623			2.08	15 %
Severe chronic hand eczema	S	S	S	S	R	S	R	V	S								
Relistor	0.706	0.794	0.088	EQ-5D	0.0 %	US, Canada	kr 465 000	72	Lifetime (0.33)	9.864	9.158	93 %	7.478	6.772	91 %	6.77	91 %
Opioid-induced constipation in cancer treatr	V	V	V	V	R, S	A	R	A	X								
Ebixa	1.214	1.238	0.024	EQ-5D	3.0 %	UK	kr 315 000	80	Lifetime (5)	5.580	4.366	78 %	4.824			3.61	75 %
Alzheimer dementia	R	R	R		R	A	R	S	S								
Prolia	8.048	8.067	0.019	EQ-5D	4.0 %	International	kr 283 067	71	Lifetime (100)	10.438	2.390	23 %	7.813			-0.23	-3 %
Osteoporosis	H	H	H	R	H	S	R	R									
Revimid	2.190	2.950	0.760	EQ-5D	4.0 %	International	kr 247 078	63	Lifetime	15.569	13.379	86 %	10.535	8.345	79 %	8.35	79 %
Relapse refractory multiple myeloma	R	R			R	S	R	S	S								
Onbrez	2.142	2.150	0.008	EQ-5D	4.0 %	Spain	kr 154 000	64	3	3.134	0.992	32 %	2.960	0.818	28 %	0.82	28 %
Chronic obstructive pulmonary disease (COPD)	R	R	R	R	S	R	R	S	S								
Victoza	7.180	7.400	0.220	EQ-5D	4.0 %	Denmark	kr 249 653	57	Lifetime (40)	19.752	12.572	64 %	12.338	5.158	42 %	5.16	42 %
Diabetes type 2 obese patients	R	R	R	S	S	S	R	S	S								
Targiniq	0.213	0.220	0.007	EQ-5D (SF-36)	0.0 %	Europe, Israel	kr 217 316	59	0.5	18.333	18.120	99 %	11.763	11.551	98 %	11.55	98 %
Severe pain opioid induced constipation	S	S	S	S	R	S	R	S	S								
Aldara	0.993	0.994	0.0007	TTO	0.0 %	International	kr 991 763	69	1	0.794	-0.199	-25 %	0.794	-0.199	-25 %	-0.20	-25 %
Actinic keratosis	S	S	S	A	S	S	R	S, U	S								
Firmagon	0.968	1.020	0.052	N/A	4.0 %	Norway	kr -867 237	73	Lifetime	9.286	8.318	90 %	7.127	6.159	86 %	6.16	86 %
Hormone-dependent metastatic prostate ca	R	R	R		R	S	S	R									
Glivec (CML)	5.000	9.960	4.960	EQ-5D	4.0 %	International	kr 384 114	51	Lifetime (60)	24.114	19.114	79 %	13.887	8.887	64 %	8.89	64 %
Chronic myelogenous leukemia	S	S	S	A	R	R	R	S	S								
Glivec (GIST)	2.170	4.020	1.850	EQ-5D	4.0 %	International	kr 268 400	60	Lifetime (10)	17.633	15.463	88 %	11.467	9.297	81 %	9.30	81 %
Gastrointestinal Stromal Tumors	R	R	R	A	R	S	R	A	S								
Efient	9.740	9.780	0.040	EQ-5D (US QALE)	4.0 %	International	kr 23 482	61	Lifetime (40)	16.934	7.194	42 %	11.160	1.420	13 %	1.42	13 %
Acute Coronary Syndromes	R	R	R	S	S	S	R	S	S								
Nexavar	0.690	1.120	0.430	EQ-5D	2.5 %	International	kr 679 515	67	Lifetime (14)	12.947	12.257	95 %	9.225	8.535	93 %	8.53	93 %
Hepatocellular carcinoma	V	V	V	S	R	S	R	S	S								
Mezavant	3.441	3.459	0.018	EQ-5D	3.0 %	US	kr -287 051	43	5	4.082	0.641	16 %	3.852			0.41	11 %
Ulcerative colitis	R	R	R	R	R	S	R	R	S								
Thalidomide Pharmorion	2.290	3.170	0.880	EQ-5D	4.0 %	Belgium, France	kr 256 547	79	Lifetime (29)	6.081	3.791	62 %	4.986	2.696	54 %	2.70	54 %
Multiple myeloma	S	S	S	S	R	V	R	V	S								
Aclasta	6.166	6.201	0.035	EQ-5D	4.0 %	International	kr -439 886	73	Lifetime	9.286	3.120	34 %	7.127	0.961	13 %	0.96	13 %
Osteoporosis	S	S	S	S	R	A	R	S	S								
Multaq	6.543	6.804	0.261	EQ-5D	4.0 %	US	kr 49 604	72	Lifetime (28)	9.864	3.321	34 %	7.478	0.935	13 %	0.93	13 %
Atrial fibrillation	V	V	V	R	R	S	V	V	V								

Table 16 Severity calculations